OM nucleic - nucleic search, using sw model

Run on: April 30, 2004, 02:25:13 ; Search time 6630 Seconds
(without alignments)
2588.815 Million cell updates/sec

Title: US-09-684-215B-3
Perfect score: 396
Sequence: 1 acggccgcgtccgataactt.....tggccgagggacccccggcc 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	396	100.0	702	6	BD251334	BD251334 Fused pro
2	396	100.0	702	6	AR403747	AR403747 Sequence
3	396	100.0	1068	6	BD274032	BD274032 Sequences
4	396	100.0	1068	6	AX005788	AX005788 Sequence
5	396	100.0	1143	6	BD274033	BD274033 Sequences
6	396	100.0	1143	6	AX005790	AX005790 Sequence
7	396	100.0	2287	6	BD251322	BD251322 Fused pro
8	396	100.0	2287	6	AR303127	AR303127 Sequence
9	396	100.0	2287	6	AR403735	AR403735 Sequence
10	396	100.0	14029	1	AE006925	AE006925 Mycobacte
11	396	100.0	341957	15	BX842572	Bx842572 Mycobacte
12	396	100.0	343050	1	BX248334	BX248334 Mycobacte
13	394.4	99.6	447	6	AR169152	AR169152 Sequence
14	394.4	99.6	447	6	AR182442	AR182442 Sequence
15	394.4	99.6	447	6	AR194825	AR194825 Sequence
16	394.4	99.6	447	6	AR233097	AR233097 Sequence
17	394.4	99.6	447	6	AR353302	AR353302 Sequence
18	394.4	99.6	447	6	AX429596	AX429596 Sequence
19	394.4	99.6	447	6	AX832581	AX832581 Sequence
20	394.4	99.6	447	6	BD006325	BD006325 Compounds
21	394.4	99.6	447	6	BD006445	BD006445 Compounds
22	394.4	99.6	447	6	BD069285	BD069285 Compounds
23	394.4	99.6	447	6	BD205817	BD205817 Compounds
24	391.2	98.8	1872	6	AR169165	AR169165 Sequence
25	391.2	98.8	1872	6	AR182455	AR182455 Sequence
26	391.2	98.8	1872	6	AR194838	AR194838 Sequence
27	391.2	98.8	1872	6	AR233110	AR233110 Sequence
28	391.2	98.8	1872	6	AR353315	AR353315 Sequence
29	391.2	98.8	1872	6	AX429609	AX429609 Sequence
30	391.2	98.8	1872	6	AX832594	AX832594 Sequence
31	391.2	98.8	1872	6	BD006338	BD006338 Compounds
32	391.2	98.8	1872	6	BD006458	BD006458 Compounds
33	391.2	98.8	1872	6	BD069298	BD069298 Compounds
34	391.2	98.8	1872	6	BD205830	BD205830 Compounds
35	354	89.4	675	6	AR261272	AR261272 Sequence
36	354	89.4	675	6	AR400535	AR400535 Sequence
37	354	89.4	675	6	AR405802	AR405802 Sequence
38	354	89.4	675	6	AX201049	AX201049 Sequence
39	354	89.4	675	6	AX267848	AX267848 Sequence
40	354	89.4	822	6	AR277645	AR277645 Sequence
41	354	89.4	822	6	AX369152	AX369152 Sequence
42	354	89.4	894	6	AX351489	AX351489 Sequence
43	354	89.4	900	6	AR220690	AR220690 Sequence
44	354	89.4	900	6	AX365960	AX365960 Sequence
45	354	89.4	915	6	AR261281	AR261281 Sequence

ALIGNMENTS

RESULT 1
BD251334
LOCUS BD251334 702 bp DNA linear PAT 17-JUL-2003
DEFINITION Fused protein of Mycobacterium tuberculosis antigen and utilization thereof.
ACCESSION BD251334
VERSION BD251334.1 GI:33061104
KEYWORDS JP 2002510494-A/13.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 702)
AUTHORS Skeiky,Y.A.W., Alderson,M. and Neto,A.C.
TITLE Fused protein of Mycobacterium tuberculosis antigen and utilization thereof

```
  JOURNAL   Patent: JP 2002510494-A 13 09-APR-2002;
            CORIXA CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002510494-A/13
            PD   09-APR-2002
            PF   07-APR-1999 JP 2000542460
            PR   07-APR-1998 US    09/056556,30-DEC-1998 US    09/223040 PI
            YASIR A W SKEIKY,MARK ALDERSON,ANTONIO CAMPOS NETO PC
            C12N15/09,A61K39/04,A61K48/00,A61P31/04,C07K14/35,C07K19/00, PC
            C12P21/02,
            PC   C12N15/00
            CC   Description of Artificial Sequence:bi-fusion
            protein Ra12-DPPD
            CC            (designated
            CC   Mtb24), reading frame 1
            CC   bi-fusion protein Ra12-DPPD (designated Mtb24),reading frame 1
            CC   reading frame 2
            CC   reading frame 3
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(696)
            FT   CDS             (2). .(700)
            FT   CDS             (3). .(701).
FEATURES              Location/Qualifiers
     source           1. .702
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 396;  DB 6;  Length 702;
  Best Local Similarity  100.0%;  Pred. No. 4.5e-57;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 2
AR403747
LOCUS       AR403747                  702 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 27 from patent US 6627198.
ACCESSION   AR403747
VERSION     AR403747.1  GI:40151423
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of Mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6627198-A 27 30-SEP-2003;
FEATURES              Location/Qualifiers
     source           1. .702
                      /organism="unknown"
                      /mol_type="genomic DNA"
ORIGIN

  Query Match           100.0%;  Score 396;  DB 6;  Length 702;
  Best Local Similarity  100.0%;  Pred. No. 4.5e-57;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420


RESULT 3
BD274032
LOCUS       BD274032                  1068 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose.
ACCESSION   BD274032
VERSION     BD274032.1  GI:33083800
KEYWORDS    JP 2002534956-A/256.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Gicquel,B., Portnoi,D., Lim,E., Pelicic,V., Guigueno,A. and
            Salmoniere,Y.G.D.L.
  TITLE     Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose
  JOURNAL   Patent: JP 2002534956-A 256 22-OCT-2002;
            INSTITUT PASTEUR
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002534956-A/256
            PD   22-OCT-2002
            PF   14-AUG-1998 JP 2000509849
            PR   14-AUG-1997 FR     97/10404,11-SEP-1997 FR     97/11325 PI
            BRIGITTE GICQUEL,DENIS PORTNOI,ENG-MONG LIM,VLADIMIR PELICIC, PI
            AGNES GUIGUENO,YVES GOGUET DE LA SALMONIERE
            PC   C12N15/09,A61K39/04,A61K48/00,A61P11/00,A61P31/04,C07K14/35,
```

```
            PC   C07K16/12,
            PC   C07K19/00,C12N1/21,C12P21/02,C12Q1/02,C12Q1/68,G01N33/50, PC
            G01N33/53//
            PC   (C12N15/09,C12R1:32),C12N15/00,(C12N15/00,C12R1:32) CC   50D
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(1065).
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1773"
ORIGIN

  Query Match           100.0%;  Score 396;  DB 6;  Length 1068;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-57;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065


RESULT 4
AX005788
LOCUS       AX005788               1068 bp    DNA     linear   PAT 24-AUG-2000
DEFINITION  Sequence 907 from Patent WO9909186.
ACCESSION   AX005788
VERSION     AX005788.1  GI:9928795
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Portnoi,D. and Guigueno,A.
  TITLE     Polypeptide nucleic sequences exported from mycobacteria, vectors
            comprising same and uses for diagnosing and preventing tuberculosis
  JOURNAL   Patent: WO 9909186-A 907 25-FEB-1999;
            PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1773"
     CDS             1. .1068
                     /note="unnamed protein product; 50D"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC05178.1"
                     /db_xref="GI:9928796"
                     /db_xref="REMTREMBL:CAC05178"
                     /translation="MSNSRRRSLRWSWLLSVLAAVGLGLATAPAQAAPPALSQDRFAD
                     FPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGAGTGIVIDPNGVVLTNNHVIAGATD
                     INAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSG
                     GQGGTPRAVPGRVVALGQTVQASDSLTGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQ
                     VVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVVD
                     NNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVT
                     WQTKSGGTRTGNVTLAEGPPA"
ORIGIN

  Query Match           100.0%;  Score 396;  DB 6;  Length 1068;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-57;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065


RESULT 5
BD274033
LOCUS       BD274033               1143 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose.
ACCESSION   BD274033
VERSION     BD274033.1  GI:33083801
KEYWORDS    JP 2002534956-A/257.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 1143)
  AUTHORS   Gicquel,B., Portnoi,D., Lim,E., Pelicic,V., Guigueno,A. and
            Salmoniere,Y.G.D.L.
  TITLE     Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose
  JOURNAL   Patent: JP 2002534956-A 257 22-OCT-2002;
            INSTITUT PASTEUR
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002534956-A/257
            PD   22-OCT-2002
            PF   14-AUG-1998 JP 2000509849
            PR   14-AUG-1997 FR     97/10404,11-SEP-1997 FR     97/11325 PI
            BRIGITTE GICQUEL,DENIS PORTNOI,ENG-MONG LIM,VLADIMIR PELICIC, PI
```

AGNES GUIGUENO,YVES GOGUET DE LA SALMONIERE
PC C12N15/09,A61K39/04,A61K48/00,A61P11/00,A61P31/04,C07K14/35,
PC C07K16/12,
PC C07K19/00,C12N1/21,C12P21/02,C12Q1/02,C12Q1/68,G01N33/50, PC
G01N33/53//
PC (C12N15/09,C12R1:32),C12N15/00,(C12N15/00,C12R1:32) CC SEQ
ID No. 50F
FH Key Location/Qualifiers
FT CDS (4). .(1140).

FEATURES Location/Qualifiers
source 1. .1143
/organism="Mycobacterium tuberculosis"
/mol_type="genomic DNA"
/db_xref="taxon:1773"

ORIGIN

Query Match 100.0%; Score 396; DB 6; Length 1143;
Best Local Similarity 100.0%; Pred. No. 4.1e-57;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1140
```

RESULT 6
AX005790
LOCUS AX005790 1143 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 909 from Patent WO9909186.
ACCESSION AX005790
VERSION AX005790.1 GI:9928797
KEYWORDS .
SOURCE Mycobacterium tuberculosis
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1
AUTHORS Portnoi,D. and Guigueno,A.
TITLE Polypeptide nucleic sequences exported from mycobacteria, vectors
comprising same and uses for diagnosing and preventing tuberculosis
JOURNAL Patent: WO 9909186-A 909 25-FEB-1999;
PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES Location/Qualifiers
source 1. .1143
/organism="Mycobacterium tuberculosis"
/mol_type="unassigned DNA"
/db_xref="taxon:1773"
CDS <4. .1143
/note="unnamed protein product; SEQ ID NO 50F"
/codon_start=1
/transl_table=11
/protein_id="CAC05179.1"
/db_xref="GI:9928798"
/db_xref="REMTREMBL:CAC05179"
/translation="KNPAARTLKAGTISDSYPDTGGYGMSNSRRRSLRWSWLLSVLAA
VGLGLATAPAQAAPPALSQDRFADFPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGA
GTGIVIDPNGVVLTNNHVIAGATDINAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAG
GLPSAAIGGGVAVGEPVVAMGNSGGQGGTPRAVPGRVVALGQTVQASDSLTGAEETLN
GLIQFDAAIQPGDSGGPVVNGLGQVVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQ
IRSGGGSPTVHIGPTAFLGLGVVDNNGNGARVQRVVGSAPAASLGISTGDVITAVDGA
PINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNVTLAEGPPA"

ORIGIN

Query Match 100.0%; Score 396; DB 6; Length 1143;
Best Local Similarity 100.0%; Pred. No. 4.1e-57;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1140
```

RESULT 7
BD251322
LOCUS BD251322 2287 bp DNA linear PAT 17-JUL-2003
DEFINITION Fused protein of Mycobacterium tuberculosis antigen and utilization
thereof.
ACCESSION BD251322
VERSION BD251322.1 GI:33061092
KEYWORDS JP 2002510494-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 2287)
AUTHORS Skeiky,Y.A.W., Alderson,M. and Neto,A.C.
TITLE Fused protein of Mycobacterium tuberculosis antigen and utilization
thereof
JOURNAL Patent: JP 2002510494-A 1 09-APR-2002;
CORIXA CORP
COMMENT OS Artificial Sequence
PN JP 2002510494-A/1
PD 09-APR-2002
PF 07-APR-1999 JP 2000542460
PR 07-APR-1998 US 09/056556,30-DEC-1998 US 09/223040 PI
YASIR A W SKEIKY,MARK ALDERSON,ANTONIO CAMPOS NETO PC
C12N15/09,A61K39/04,A61K48/00,A61P31/04,C07K14/35,C07K19/00, PC
C12P21/02,

```
            PC   C12N15/00
            CC   Description of Artificial Sequence:tri-fusion protein Ra12- CC
                 TbH9-Ra35
            CC   (designated Mtb32A)
            CC   n = g, a, c or t
            CC   n = g, a, c or t
            CC   n = g, a, c or t
            FH   Key             Location/Qualifiers
            FT   modified_base   (30)
            FT   modified_base   (33)
            FT   CDS             (42). .(2231)
            FT   modified_base   (2270).
FEATURES            Location/Qualifiers
     source         1. .2287
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
ORIGIN
```

Query Match 100.0%; Score 396; DB 6; Length 2287;
Best Local Similarity 100.0%; Pred. No. 3.6e-57;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

```
RESULT 8
AR303127
LOCUS       AR303127                2287 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 1 from patent US 6544522.
ACCESSION   AR303127
VERSION     AR303127.1  GI:31691855
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Skeiky,Y., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6544522-A 1 08-APR-2003;
FEATURES            Location/Qualifiers
     source         1. .2287
                    /organism="unknown"
                    /mol_type="genomic DNA"
ORIGIN
```

Query Match 100.0%; Score 396; DB 6; Length 2287;
Best Local Similarity 100.0%; Pred. No. 3.6e-57;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

```
RESULT 9
AR403735
LOCUS       AR403735                2287 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1 from patent US 6627198.
ACCESSION   AR403735
VERSION     AR403735.1  GI:40151411
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of Mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6627198-A 1 30-SEP-2003;
FEATURES            Location/Qualifiers
     source         1. .2287
                    /organism="unknown"
                    /mol_type="genomic DNA"
ORIGIN
```

Query Match 100.0%; Score 396; DB 6; Length 2287;
Best Local Similarity 100.0%; Pred. No. 3.6e-57;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302
```

```
Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

RESULT 10
AE006925
LOCUS AE006925 14029 bp DNA linear BCT 27-APR-2001
DEFINITION Mycobacterium tuberculosis CDC1551, section 11 of 280 of the complete genome.
ACCESSION AE006925 AE000516
VERSION AE006925.1 GI:13879610
KEYWORDS .
SOURCE Mycobacterium tuberculosis CDC1551
ORGANISM Mycobacterium tuberculosis CDC1551
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 14029)
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O., Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E., Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M., Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H., Gill,J., Mikula,A. and Bishai,W.
TITLE Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 14029)
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O., Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E., Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M., Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H., Gill,J., Mikula,A. and Bishai,W.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES Location/Qualifiers

```
     source          1. .14029
                     /organism="Mycobacterium tuberculosis CDC1551"
                     /mol_type="genomic DNA"
                     /strain="CDC1551"
                     /db_xref="taxon:83331"
                     /note="clinical strain"
     gene            complement(99. .533)
                     /gene="MT0129"
     CDS             complement(99. .533)
                     /gene="MT0129"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK44353.1"
                     /db_xref="GI:13879611"
                     /translation="MGEFDPKLRFAQSPVARLATSTPDGTPHLVPVVFALGARRPAEA
                     TGADVIYTAVDAKRKTTQRLRRLANLEHNPRASVLVDSYADDWTQLWWVRADGVAAIH
                     RDGEVMRAAYRLLRAKYTQYQSVPLNGPVIAIAVQRWASRHA"
     gene            complement(643. .1062)
                     /gene="MT0130"
     CDS             complement(643. .1062)
                     /gene="MT0130"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK44354.1"
                     /db_xref="GI:13879612"
                     /translation="MASSSSLPLLRHCVSDDQVTVVGFDGDDLGKTARRIAALVVQRA
                     IFLNDRNTAVAHSGDDAVLGHAVLPGVPRDPDPLHASSMYSILGMCQSVNGRPFDAIA
                     LVSVRLCHVQTDPTDSCGGRDRPGQLPCAPLDYHRHH"
     gene            1047. .1415
                     /gene="MT0131"
     CDS             1047. .1415
                     /gene="MT0131"
                     /note="identified by match to protein family HMM"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA-binding protein, CopG family"
                     /protein_id="AAK44355.1"
                     /db_xref="GI:13879613"
                     /translation="MTKKPRNPADYVIGDDVEVSDVDLKQEEVYVDGERLTDERVEQM
                     ASESLRLAREREANLIPGGKSLSGGSAHSPAVQVVVSKATHAKLKELARSRKMSVSKL
                     LRPVLDEFVQRETGRILPRR"
     gene            1682. .3367
                     /gene="MT0132"
     CDS             1682. .3367
                     /gene="MT0132"
                     /note="similar to GB:AL123456; identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="PE_PGRS family protein"
                     /protein_id="AAK44356.1"
                     /db_xref="GI:13879614"
                     /translation="MFAGGGAGGLGRCVMSFVSVAPEIVVAAATDLAGIGSAISAANA
                     AAAAPTTAVLAAGADEVSAAIAALFSGHAQAYQALSAQAAAFHQQFVQTLAGGAGAYA
                     AAEAQVEQQLLAAINAPTQALLGRPLIGNGADGAPGTGQAGGAGGILYGNGGNGGSGA
                     AGQAGGAGGPAGLIGHGGSGGAGGSGAAGGAGGHGGWLWGNGGVGGSGGAGVGAGVAG
                     GHGGAGGAAGLWGAGGGGGNGGNGADANIVSGGDGGLGGAGGGGGWLYGDGGAGGHGG
                     QGAGGGAGGAGGDGGQGGAGRGLWGTGGAGGHGGQGGGTGGPPLPGQAGMGAAGGAGG
                     LIGNGGAGGDGGVGASGGVAGVXGAGGNAMLIGHGGAGGAGGDSSFANGAAGGAGGAG
                     GHLFGNGGSGGHGGAVTAGNTGIGGAGGVGGDARLIGHGGAGGAGGDRAGALVGRDGG
                     PGGNGGAGGQLYGNGGDGGPGGQGGQAFGANNIGGTGGAGGNGGPAILSGNGGNGGAG
                     GAGGAGGAGGGAGGVGGAGGAPGTGGTLQAAVSGLVTALFGAPGQPGDTGQPG"
     gene            3519. .4586
                     /gene="MT0133"
     CDS             3519. .4586
                     /gene="MT0133"
                     /note="identified by match to protein family HMM"
                     /codon_start=1
                     /transl_table=11
                     /product="serine protease, putative"
                     /protein_id="AAK44357.1"
                     /db_xref="GI:13879615"
                     /translation="MSNSRRRSLRWSWLLSVLAAVGLGLATAPAQAAPPALSQDRFAD
                     FPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGAGTGIVIDPNGVVLTNNHVIAGATD
                     INAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSG
                     GQGGTPRAVPGRVVALGQTVQASDSLTGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQ
                     VVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVVD
                     NNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVT
                     WQTKSGGTRTGNVTLAEGPPA"
     gene            4695. .6500
                     /gene="MT0134"
     CDS             4695. .6500
                     /gene="MT0134"
                     /note="similar to GP:2808807; identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-amylase family protein"
                     /protein_id="AAK44358.1"
                     /db_xref="GI:13879616"
                     /translation="MNEAEHSVEHPPVQGSHVEGGVVEHPDAKDFGSAAALPADPTWF
                     KHAVFYEVLVRAFFDASADGSGDLRGLIDRLDYLQWLGIDCIWLPPFYDSPLRDGGYD
                     IRDFYKVLPEFGTVDDFVALVDAAHRRGIRIITDLVMNHTSESHPWFQESRRDPDGPY
                     GDYYVWSDTSERYTDARIIFVDTEESNWSFDPVRRQFYWHRFFSHQPDLNYDNPAVQE
                     AMIDVIRFWLGLGIDGFRLDAVPYLFEREGTNCENLPETHAFLKRVRKVVDDEFPGRV
                     LLAEANQWPGDVVEYFGDPNTGGDECHMAFHFPLMPRIFMAVRRESRFPISEIIAQTP
```

```
                PIPDMAQWGIFLRNHDELTLEMVTDEERDYMYAEYAKDPRMKANVGIRRRLAPLLDND
                RNQIELFTALLLSLPGSPVLYYGDEIGMGDVIWLGDRDGVRIPMQWTPDRNAGFSTAN
                PGRLYLPPSQDPVYGYQAVNVEAQRDTSTSLLNFTRTMLAVRRRHPAFAVGAFQELGG
                SNPSVLAYVRQVAGDDGDTVLCVNNLSRFPQPIELDLQQWTNYTPVELTGHVEFPRIG
                QVPYLLTLPGHGFYWFQLTTHEVGAPPTCGGERRL"
    gene        6603. .7970
                /gene="MT0135"
    CDS         6603. .7970
                /gene="MT0135"
                /note="similar to GP:2808801; identified by sequence
                similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="pep2 protein"
                /protein_id="AAK44359.1"
                /db_xref="GI:13879617"
                /translation="MTRSDTLATKLPWSDWLPRQRWYAGRNRELATVKPGVVVALRHN
                LDLVLVDVTYTDGATERYQVLVGWDFEPASEYGTKAAIGVADDRTGFDALYDVAGPQF
                LLSLIVSSAVCGTSTGEVTFTREPDVELPFAAQPRVCDAEQSNTSVIFDRRAILKVFR
                RVSSGINPDIELNRVLTRAGNPHVARLLGAYQFGRPNRSPTDALAYALGMVTEYEANA
                AEGWAMATASVRDLFAEGDLYAHEVGGDFAGESYRLGEAVASVHATLADSLGTAQATF
                PVDRMLARLSSTVAVVPELREYAPTIEQQFQKLAAEAITVQRVHGDLHLGQVLRTPES
                WLLIDFEGEPGQPLDERRAPDSPLRDVAGVLRSFEYAAYGPLVDQATDKQLAARAREW
                VERNRAAFCDGYAVASGIDPRDSALLLGAYELDKAVYETGYETRHRPGWLPIPLRSIA
                RLTAS"
    gene        8038. .8817
                /gene="MT0136"
    CDS         8038. .8817
                /gene="MT0136"
                /note="identified by Glimmer2; putative"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAK44360.1"
                /db_xref="GI:13879618"
                /translation="MQREIYDGEARLSWVLAALAGILGATAFTHSAGYFVTFMTGNSQ
                RAVLGLFGDDAWMSVTASLLILFFVAGVVIASVCRRHFWAAHPHGPTVLTTFSLIFAA
                GVDIMLGGWHESMLDFVPILFVVFGIGALNTSFVKDGEVSVPLSYVTGTLVKMGQGIE
                RHLAGGKVEDWLGYFLLHASFVLGAAAGGAISMVVTGPQMLAVAAVVCAATTGYTYLH
                ADRRGLVNQKRPQPGKRLFRALRRGELDSGTSTPATNYGSS"
    gene        complement(8949. .9989)
                /gene="MT0137"
    CDS         complement(8949. .9989)
                /gene="MT0137"
                /note="similar to SP:P31953 GB:X57229 PID:48828
                PID:1567734 PID:1877254; identified by sequence
                similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="esterase, putative, antigen 85-C"
                /protein_id="AAK44361.1"
                /db_xref="GI:13879619"
                /translation="MGLVVAMTFFEQVRRLRSAATTLPRRLAIAAMGAVLVYGLVGTF
                GGPATAGAFSRPGLPVEYLQVPSASMGRDIKVQFQGGGPHAVYLLDGLRAQDDYNGWD
                INTPAFEEYYQSGLSVIMPVGGQSSFYTDWYQPSQSNGQNYTYKWETFLTREMPAWLQ
                ANKGVSPTGNAAVGLSMSGGSALILAAYYPQQFPYAASLSGFLNPSEGWWPTLIGLAM
                NDSGGYNANSMWGPSSDPAWKRNDPMVQIPRLVANNTRIWVYCGNGTPSDLGGDNIPA
                KFLEGLTLRTNQTFRDTYAADGGRNGVFNFPPNGTHSWPYWNEQLVAMKADIQHVLNG
                ATPPAAPAAPAA"
    gene        10218. .10673
                /gene="MT0138"
    CDS         10218. .10673
                /gene="MT0138"
                /note="similar to SP:P08634 PID:46222; identified by
                sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="nodulation protein N-related protein"
                /protein_id="AAK44362.1"
                /db_xref="GI:13879620"

Query Match             100.0%;  Score 396;  DB 1;  Length 14029;
Best Local Similarity   100.0%;  Pred. No. 2.5e-57;
```

```
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4188 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 4247

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4248 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 4307

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4308 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 4367

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4368 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 4427

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4428 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 4487

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4488 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 4547

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       4548 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 4583


RESULT 11
BX842572
ID   BX842572   standard; circular genomic DNA; PRO; 341957 BP.
XX
AC   BX842572; AL021427; AL021428; AL021926; AL021927; AL021928; AL021929;
AC   AL021930; AL123456; Z74410; Z80233; Z80775; Z86089; Z92669; Z92770; Z96071;
AC   Z97050;
XX
SV   BX842572.1
XX
DT   21-NOV-2003 (Rel. 77, Created)
DT   21-NOV-2003 (Rel. 77, Last updated, Version 1)
XX
DE   Mycobacterium tuberculosis H37Rv complete genome; segment 1/13
XX
KW   complete genome.
XX
OS   Mycobacterium tuberculosis H37Rv
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex; Mycobacterium tuberculosis.
XX
RN   [1]
RX   MEDLINE; 98295987.
RX   PUBMED; 9634230.
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry III C.E., Tekaia F., Badcock K.,
RA   Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K.,
RA   Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K.,
RA   Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J.,
RA   Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S.,
RA   Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the complete
RT   genome sequence";
RL   Nature 393:537-544(1998).
XX
RN   [2]
RX   PUBMED; 12368430.
RA   Camus J.C., Pryor M.J., Medigue C., Cole S.T.;
RT   "Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv";
RL   Microbiology 148:2967-2973(2002).
XX
```

```
RN   [3]
RP   1-341957
RA   Parkhill J.;
RT   ;
RL   Submitted (11-JUN-1998) to the EMBL/GenBank/DDBJ databases.
RL   Submitted on behalf of the Mycobacterium tuberculosis sequencing and
RL   mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton,
RL   Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut
RL   Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail:
RL   parkhill@sanger.ac.uk
XX
CC   Notes:
CC   Details of M. tuberculosis sequencing at the Sanger Centre
CC   are available on the World Wide Web.
CC   (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/)
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .341957
FT                   /db_xref="taxon:83332"
FT                   /mol_type="genomic DNA"
FT                   /organism="Mycobacterium tuberculosis H37Rv"
FT                   /strain="H37Rv"
FT   CDS             1. .1524
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0001, (MT0001, MTV029.01, P49993), len: 507 aa.
FT                   dnaA, chromosomal replication initiator protein (see
FT                   citations below), equivalent to other Mycobacterial
FT                   CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g.
FT                   P46388|DNAA_MYCLE from Mycobacterium leprae (502 aa);
FT                   Q9L7L7|DNAA_MYCPA from Mycobacterium paratuberculosis (509
FT                   aa); P49990|DNAA_MYCAV from Mycobacterium avium (508 aa);
FT                   P49992|DNAA_MYCSM from Mycobacterium smegmatis (504 aa);
FT                   etc. Also highly similar to others except in N-terminus
FT                   e.g. Q9ZH75|DNAA_STRCH CHROMOSOMAL REPLICATION INITIATOR
FT                   PROTEIN from Streptomyces chrysomallus (624 aa);
FT                   Q9ZH76|DNAA_STRRE from Streptomyces reticuli (643 aa);
FT                   DNAA_ECOLI|P03004|B3702 chromosomal replication initiator
FT                   protein from Escherichia coli strain K12 (467 aa), FASTA
FT                   scores: opt: 986, E(): 0, (43.2% identity in 389 aa
FT                   overlap); etc. Contains PS00017 ATP/GTP-binding site motif
FT                   A (P-loop) and PS01008 DnaA protein signature. BELONGS TO
FT                   THE DNAA FAMILY. Note that the first base of this gene has
FT                   been taken as base 1 of the Mycobacterium tuberculosis
FT                   H37Rv genomic sequence."
FT                   /transl_table=11
FT                   /function="PLAYS AN IMPORTANT ROLE IN THE INITIATION AND
FT                   REGULATION OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN
FT                   OF REPLICATION; IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA
FT                   AT A 9 BP CONSENSUS (DNAA BOX): 5'-TTATC(C/A)A(C/A)A-3'.
FT                   DNAA BINDS TO ATP AND TO ACIDIC PHOSPHOLIPIDS. DNAA PROTEIN
FT                   BINDS THE ORIGIN OF REPLICATION (oriC), ATP AND ADP, AND
FT                   EXHIBITED WEAK ATPase ACTIVITY."
FT                   /gene="dnaA"
FT                   /locus_tag="Rv0001"
FT                   /product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
FT                   /protein_id="CAA16238.1"
FT                   /translation="MTDDPGSGFTTVWNAVVSELNGDPKVDDGPSSDANLSAPLTPQQR
FT                   AWLNLVQPLTIVEGFALLSVPSSFVQNEIERHLRAPITDALSRRLGHQIQLGVRIAPPA
FT                   TDEADDTTVPPSENPATTSPDTTTDNDEIDDSAAARGDNQHSWPSYFTERPHNTDSATA
FT                   GVTSLNRRYTFDTFVIGASNRFAHAAALAIAEAPARAYNPLFIWGESGLGKTHLLHAAG
FT                   NYAQRLFPGMRVKYVSTEEFTNDFINSLRDDRKVAFKRSYRDVDVLLVDDIQFIEGKEG
FT                   IQEEFFHTFNTLHNANKQIVISSDRPPKQLATLEDRLRTRFEWGLITDVQPPELETRIA
FT                   ILRKKAQMERLAVPDDVLELIASSIERNIRELEGALIRVTAFASLNKTPIDKALAEIVL
FT                   RDLIADANTMQISAATIMAATAEYFDTTVEELRGPGKTRALAQSRQIAMYLCRELTDLS
FT                   LPKIGQAFGRDHTTVMYAQRKILSEMAERREVFDHVKELTTRIRQRSKR"
FT   misc_feature    622. .645
FT                   /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
FT   misc_feature    1384. .1440
FT                   /note="PS01008 DnaA protein signature"
FT   CDS             2052. .3260
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0002, (MTV029.02, MTCY10H4.0), len: 402 aa. dnaN,
FT                   DNA polymerase III (beta chain) (EC 2.7.7.7) (see citations
FT                   below), equivalent to other Mycobacterial DNA POLYMERASES
FT                   III BETA CHAIN e.g. NP_301130.1|NC_002677 from
FT                   Mycobacterium leprae (399 aa); Q9L7L6|DP3B_MYCPA from
FT                   Mycobacterium avium subsp. paratuberculosis (399 aa);
FT                   P52851|DP3B_MYCSM from Mycobacterium smegmatis (397 aa);
FT                   etc. Also highly similar to others e.g. P27903|DP3B_STRCO
FT                   DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor
FT                   (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity
FT                   in 337 aa overlap); P21174|DP3B_MICLU from Micrococcus
FT                   luteus (310 aa); P52023|DP3B_SYNP7 from Synechococcus sp.
FT                   strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in
FT                   neighbouring cosmid MTCY10H4.01."
FT                   /transl_table=11
FT                   /EC_number="2.7.7.7"
FT                   /function="DNA POLYMERASE III IS A COMPLEX, MULTICHAIN
FT                   ENZYME RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN
FT                   BACTERIA. THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5'
FT                   EXONUCLEASE ACTIVITY. THE BETA CHAIN IS REQUIRED FOR
FT                   INITIATION OF REPLICATION ONCE IT IS CLAMPED ONTO DNA, IT
FT                   SLIDES FREELY (BIDIRECTIONAL AND ATP-INDEPENDENT) ALONG
FT                   DUPLEX DNA {CATALYTIC ACTIVITY: N deoxynucleoside
FT                   triphosphate = N diphosphate + {DNA}N}."
FT                   /gene="dnaN"
FT                   /locus_tag="Rv0002"
FT                   /product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA
FT                   NUCLEOTIDYLTRANSFERASE)"
FT                   /protein_id="CAA16239.1"
FT                   /translation="MDAATTRVGLTDLTFRLLRESFADAVSWVAKNLPARPAVPVLSGV
FT                   LLTGSDNGLTISGFDYEVSAEAQVGAEIVSPGSVLVSGRLLSDITRALPNKPVDVHVEG
FT                   NRVALTCGNARFSLPTMPVEDYPTLPTLPEETGLLPAELFAEAISQVAIAAGRDDTLPM
FT                   LTGIRVEILGETVVLAATDRFRLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAGIGGS
FT                   DVRLSLGTGPGVGKDGLLGISGNGKRSTTRLLDAEFPKFRQLLPTEHTAVATMDVAELI
FT                   EAIKLVALVADRGAQVRMEFADGSVRLSAGADDVGRAEEDLVVDYAGEPLTIAFNPTYL
FT                   TDGLSSLRSERVSFGFTTAGKPALLRPVSGDDRPVAGLNGNGPFPAVSTDYVYLLMPVR
FT                   LPG"
FT   CDS             3280. .4437
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0003, (MTCY10H4.01), len: 385 aa. recF, DNA
FT                   replication and repair protein (see citations below),
FT                   equivalent to others Mycobacterial DNA replication and
FT                   repair proteins e.g. NP_301131.1|NC_002677 from
FT                   Mycobacterium leprae (385 aa); Q9L7L5|RECF_MYCPA from
FT                   Mycobacterium avium subsp. paratuberculosis (385 aa);
FT                   P50916|RECF_MYCSM from Mycobacterium smegmatis (384 aa);
FT                   etc. Also highly similar to others e.g. P36176|RECF_STRCO
FT                   DNA REPLICATION AND REPAIR PROTEIN from Streptomyces
FT                   coelicolor (373 aa); NP_440892.1|NC_000911 from
FT                   Synechocystis sp. strain PCC 6803 (384 aa);
FT                   NP_469352.1|NC_003212 from Listeria innocua (370 aa); etc.
FT                   Contains PS00017 ATP/GTP-binding site motif A (P-loop),
FT                   PS00617 RecF protein signature 1, and PS00618 RecF protein
FT                   signature 2. BELONGS TO THE RECF FAMILY."
FT                   /transl_table=11
FT                   /function="THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM
FT                   AND RECOMBINATION; IT IS REQUIRED FOR DNA REPLICATION AND
FT                   NORMAL SOS INDUCIBILITY. RECF BINDS PREFERENTIALLY TO
FT                   SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS TO BIND ATP."
FT                   /gene="recF"
FT                   /locus_tag="Rv0003"
FT                   /product="DNA REPLICATION AND REPAIR PROTEIN RECF
FT                   (SINGLE-STRAND DNA BINDING PROTEIN)"
FT                   /protein_id="CAB02424.1"
FT                   /translation="MYVRHLGLRDFRSWACVDLELHPGRTVPVGPNGYGKTNLIEALWY
FT                   STTLGSHRVSADLPLIRVGTDRAVISTIVVNDGRECAVDLEIATGRVNKARLNRSSVRS
FT                   TRDVVGVLRAVLFAPEDLGLVRGDPADRRRYLDDLAIVRRPAIAAVRAEYERVLRQRTA
FT                   LLKSVPGARYRGDRGVFDTLEVWDSRLAEHGAELVAARIDLVNQLAPEVKKAYQLLAPE
FT                   SRSASIGYRASMDVTGPSEQSDIDRQLLAARLLAALAARRDAELERGVCLVGPHRDDLI
FT                   LRLGDQPAKGFASHGEAWSLAVALRLAAYQLLRVDGGEPVLLLDDVFAELDVMRRRALA
FT                   TAAESAEQVLVTAAVLEDIPAGWDARRVHIDVRADDTGSMSVVLP"
FT   misc_feature    3367. .3390
FT                   /note="PS00017 ATP/GTP-binding site motif A"
FT   misc_feature    3634. .3690
```

```
FT                   /note="PS00617 RecF protein signature 1"
FT   misc_feature    4243. .4296
FT                   /note="PS00618 RecF protein signature 2"
FT   CDS             4434. .4997
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0004, (MTCY10H4.02), len: 187 aa. Conserved
FT                   hypothetical protein (see Salazar et al., 1996), highly
FT                   similar, but longer 21 aa in N-terminus, to
FT                   AAF33696.1|AF222789 unknown protein from Mycobacterium
FT                   avium subsp. paratuberculosis (166 aa); and highly similar
FT                   to NP_301132.1|NC_002677 conserved hypothetical protein
FT                   from Mycobacterium leprae (189 aa); S70990 hypothetical
FT                   protein from Mycobacterium smegmatis (194 aa). Also highly
FT                   similar, except in N-terminal part, to

  Query Match           100.0%;  Score 396;  DB 15;  Length 341957;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-57;
  Matches  396;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151817 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 151876

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151877 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 151936

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151937 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 151996

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151997 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152056

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152057 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152116

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152117 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152176

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db     152177 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152212


RESULT 12
BX248334
LOCUS       BX248334               343050 bp    DNA     linear   BCT 11-JUN-2003
DEFINITION  Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment
            1/14.
ACCESSION   BX248334 BX248333
VERSION     BX248334.1  GI:31616762
KEYWORDS    complete genome.
SOURCE      Mycobacterium bovis subsp. bovis AF2122/97
  ORGANISM  Mycobacterium bovis subsp. bovis AF2122/97
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Garnier,T., Eiglmeier,K., Camus,J.-C., Medina,N., Mansoor,H.,
            Pryor,M., Duthoy,S., Grondin,S., Lacroix,C., Monsempe,C., Simon,S.,
            Harris,B., Atkin,R., Doggett,J., Mayes,R., Keating,L.,
            Wheeler,P.R., Parkhill,J., Barrell,B.G., Cole,S.T., Gordon,S.V. and
            Hewinson,G.
  TITLE     The complete genome sequence of Mycobacterium bovis
  JOURNAL   Online Publication
  REMARK    PNAS 10.1073/pnas.1130426100 ( Microbiology )
REFERENCE   2  (bases 1 to 343050)
  AUTHORS   Garnier,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire
            Bacterienne Institut Pasteur 28,rue du Dr Roux 75724 PARIS cedex
            15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the
            Mycobacterium bovis sequencing teams, TB Research Group, Veterinary
            Laboratories Agency Weybridge, Woodham Lane, New Haw,Addlestone,
            Surrey KT15 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton,Cambridge CB10 1SA, UK. PT4 Annotation, Genopole, Institut
            Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France.
            Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28
            rue du Docteur Roux, 75724 Paris Cedex 15, France
FEATURES             Location/Qualifiers
     source          1. .343050
                     /organism="Mycobacterium bovis subsp. bovis AF2122/97"
                     /mol_type="genomic DNA"
                     /strain="AF2122/97"
                     /db_xref="taxon:233413"
     gene            1. .1524
                     /gene="dnaA"
                     /locus_tag="Mb0001"
     CDS             1. .1524
                     /gene="dnaA"
                     /locus_tag="Mb0001"
                     /note="Mb0001, dnaA, len: 507 aa. Equivalent to Rv0001,
                     len: 507 aa, from Mycobacterium tuberculosis strain H37Rv,
                     (99.6% identity in 507 aa overlap). dnaA, chromosomal
                     replication initiator protein (see citations below),
                     equivalent to other Mycobacterial CHROMOSOMAL REPLICATION
                     INITIATOR PROTEINS e.g. P46388|DNAA_MYCLE from
                     Mycobacterium leprae (502 aa); Q9L7L7|DNAA_MYCPA from
                     Mycobacterium paratuberculosis (509 aa); P49990|DNAA_MYCAV
                     from Mycobacterium avium (508 aa); P49992|DNAA_MYCSM from
                     Mycobacterium smegmatis (504 aa); etc. Also highly similar
                     to others except in N-terminus e.g. Q9ZH75|DNAA_STRCH
                     CHROMOSOMAL REPLICATION INITIATOR PROTEIN from
                     Streptomyces chrysomallus (624 aa); Q9ZH76|DNAA_STRRE from
                     Streptomyces reticuli (643 aa); DNAA_ECOLI|P03004|B3702
                     chromosomal replication initiator protein from Escherichia
                     coli strain K12 (467 aa), FASTA scores: opt: 986, E(): 0,
                     (43.2% identity in 389 aa overlap); etc. Contains PS00017
                     ATP/GTP-binding site motif A (P-loop) and PS01008 DnaA
                     protein signature. BELONGS TO THE DNAA FAMILY. Note that
                     the first base of this gene has been taken as base 1 of
                     the Mycobacterium bovis genomic sequence."
                     /codon_start=1
                     /evidence=experimental
                     /transl_table=11
                     /product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
                     /protein_id="CAD92863.1"
                     /db_xref="GI:31616763"
                     /translation="MTDDPGSGFTTVWNAVVSELNGDPKVDDGPSSDANLSAPLTPQQ
                     RAWLNLVQPLTIVEGFALLSVPSSFVQNEIERHLRAPITDALSRRLGHQIQLGVRIAP
                     PATDEADDTTVPPSENPATTSPDTTTDNDEIDDSAAARGDNQHSWPSYFTERPRNTDS
                     ATAGVTSLNRRYTFDTFVIGASNRFAHAAALAIAEAPARAYNPLFIWGESGLGKTHLL
                     HAAGNYAQRLFPGMRVKYVSTEEFTNDFINSLRDDRKVAFKRSYRDVDVLLVDDIQFI
                     EGKEGIQEEFFHTFNTLHNANKQIVISSDRPPKQLATLEDRLRTRFEWGLITDVQPPE
                     LETRIAILRKKAQMERLAIPDDVLELIASSIERNIRELEGALIRVTAFASLNKTPIDK
                     ALAEIVLRDLIADANTMQISAATIMAATAEYFDTTVEELRGPGKTRALAQSRQIAMYL
                     CRELTDLSLPKIGQAFGRDHTTVMYAQRKILSEMAERREVFDHVKELTTRIRQRSKR"
     gene            2052. .3260
                     /gene="dnaN"
                     /locus_tag="Mb0002"
     CDS             2052. .3260
                     /gene="dnaN"
                     /locus_tag="Mb0002"
                     /EC_number="2.7.7.7"
                     /note="Mb0002, dnaN, len: 402 aa. Equivalent to Rv0002,
                     len: 402 aa, from Mycobacterium tuberculosis strain H37Rv,
                     (99.8% identity in 402 aa overlap). dnaN, DNA polymerase
                     III (beta chain) (EC 2.7.7.7) (see citations below),
                     equivalent to other Mycobacterial DNA POLYMERASES III BETA
                     CHAIN e.g. NP_301130.1|NC_002677 from Mycobacterium leprae
                     (399 aa); Q9L7L6|DP3B_MYCPA from Mycobacterium avium
                     subsp. paratuberculosis (399 aa); P52851|DP3B_MYCSM from
```

Mycobacterium smegmatis (397 aa); etc. Also highly similar to others e.g. P27903|DP3B_STRCO DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity in 337 aa overlap); P21174|DP3B_MICLU from Micrococcus luteus (310 aa); P52023|DP3B_SYNP7 from Synechococcus sp. strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in neighbouring cosmid MTCY10H4.01."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA NUCLEOTIDYLTRANSFERASE)"
/protein_id="CAD92864.1"
/db_xref="GI:31616764"
/translation="MDAATTRVGLTDLTFRLLRESFADAVSWVAKNLPARPAVPVLSG VLLTGSDNGLTISGFDYEVSAEAQVGAEIVSPGSVLVSGRLLSDITRALPNKPVGVHV EGNRVALTCGNARFSLPTMPVEDYPTLPTLPEETGLLPAELFAEAISQVAIAAGRDDT LPMLTGIRVEILGETVVLAATDRFRLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAG IGGSDVRLSLGTGPGVGKDGLLGISGNGKRSTTRLLDAEFPKFRQLLPTEHTAVATMD VAELIEAIKLVALVADRGAQVRMEFADGSVRLSAGADDVGRAEEDLVVDYAGEPLTIA FNPTYLTDGLSSLRSERVSFGFTTAGKPALLRPVSGDDRPVAGLNGNGPFPAVSTDYV YLLMPVRLPG"

gene 3280. .4437
/gene="recF"
/locus_tag="Mb0003"

CDS 3280. .4437
/gene="recF"
/locus_tag="Mb0003"
/note="Mb0003, recF, len: 385 aa. Equivalent to Rv0003, len: 385 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 385 aa overlap). recF, DNA replication and repair protein (see citations below), equivalent to others Mycobacterial DNA replication and repair proteins e.g. NP_301131.1|NC_002677 from Mycobacterium leprae (385 aa); Q9L7L5|RECF_MYCPA from Mycobacterium avium subsp. paratuberculosis (385 aa); P50916|RECF_MYCSM from Mycobacterium smegmatis (384 aa); etc. Also highly similar to others e.g. P36176|RECF_STRCO DNA REPLICATION AND REPAIR PROTEIN from Streptomyces coelicolor (373 aa); NP_440892.1|NC_000911 from Synechocystis sp. strain PCC 6803 (384 aa); NP_469352.1|NC_003212 from Listeria innocua (370 aa); etc. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00617 RecF protein signature 1, and PS00618 RecF protein signature 2. BELONGS TO THE RECF FAMILY."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA REPLICATION AND REPAIR PROTEIN RECF (SINGLE-STRAND DNA BINDING PROTEIN)"
/protein_id="CAD92865.1"
/db_xref="GI:31616765"
/translation="MYVRHLGLRDFRSWACVDLELHPGRTVFVGPNGYGKTNLIEALW YSTTLGSHRVSADLPLIRVGTDRAVISTIVVNDGRECAVDLEIATGRVNKARLNRSSV RSTRDVVGVLRAVLFAPEDLGLVRGDPADRRRYLDDLAIVRRPAIAAVRAEYERVVRQ RTALLKSVPGARYRGDRGVFDTLEVWDSRLAEHGAELVAARIDLVNQLAPEVKKAYQL LAPESRSASIGYRASMDVTGPSEQSDTDRQLLAARLLAALAARRDAELERGVCLVGPH RDDLILRLGDQPAKGFASHGEAWSLAVALRLAAYQLLRVDGGEPVLLLDDVFAELDVM RRRALATAAESAEQVLVTAAVLEDIPAGWDARRVHIDVRADDTGSMSVVLP"

gene 4434. .4997
/locus_tag="Mb0004"

CDS 4434. .4997
/locus_tag="Mb0004"
/note="Mb0004, -, len: 187 aa. Equivalent to Rv0004, len: 187 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 187 aa overlap). Conserved hypothetical protein (see citation below), highly similar, but longer 21 aa in N-terminus, to AAF33696.1|AF222789 unknown protein from Mycobacterium avium subsp. paratuberculosis (166 aa); and highly similar to NP_301132.1|NC_002677 conserved hypothetical protein from Mycobacterium leprae (189 aa); S70990 hypothetical protein from Mycobacterium smegmatis (194 aa). Also similar to in C-terminus to C-terminal part of P35925|YREG_STRCO HYPOTHETICAL 19.8 KDA PROTEIN (IN RECF-GYRB INTERGENIC REGION) from Streptomyces coelicolor (190 aa), FASTA scores: opt: 404, E(): 3.9e-18, (40.7% identity in 189 aa overlap)."
/codon_start=1
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD92866.1"
/db_xref="GI:31616766"
/translation="MTGSVDRPDQNRGERLMKSPGLDLVRRTLDEARAAARARGQDAG RGRVASVASGRVAGRRRSWSGPGPDIRDPQPLGKAARELAKKRGWSVRVAEGMVLGQW SAVVGHQIAEHARPTALNDGVLSVIAESTAWATQLRIMQAQLLAKIAAAVGNDVVRSL KITGPAAPSWRKGPRHIAGRGPRDTYG"

gene 5123. .7267
/gene="gyrB"
/locus_tag="Mb0005"

CDS 5123. .7267
/gene="gyrB"
/locus_tag="Mb0005"
/EC_number="5.99.1.3"
/note="Mb0005, gyrB, len: 714 aa. Equivalent to Rv0005, len: 714 aa, from Mycobacterium tuberculosis strain H37Rv, (99.9% identity in 714 aa overlap). gyrB, DNA gyrase subunit B (EC 5.99.1.3) (see citations below), equivalent, except in N-terminus, to other Mycobacterial DNA GYRASES SUBUNIT B e.g. T10005 from Mycobacterium leprae (697 aa); Q9L7L3|GYRB_MYCPA from Mycobacterium avium subsp. paratuberculosis (677 aa) (has its N-terminus shorter); P48355|GYRB_MYCSM from Mycobacterium smegmatis (675 aa); etc. Also highly similar to others e.g. T10969 from Streptomyces coelicolor (686 aa); P50075|GYBS_STRSH from

Query Match 100.0%; Score 396; DB 1; Length 343050;
Best Local Similarity 100.0%; Pred. No. 1.4e-57;
Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152008 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 152067

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152068 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 152127

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152128 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 152187

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152188 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152247

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152248 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152307

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152308 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152367

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db     152368 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152403
```

RESULT 13
AR169152
LOCUS AR169152 447 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 4 from patent US 6290969.
ACCESSION AR169152
VERSION AR169152.1 GI:17906927
KEYWORDS .
SOURCE Unknown.

```
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6290969-A 4 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             99.6%;  Score 394.4;  DB 6;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 9e-57;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 14
AR182442
LOCUS       AR182442                 447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6338852.
ACCESSION   AR182442
VERSION     AR182442.1  GI:20225649
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6338852-A 4 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             99.6%;  Score 394.4;  DB 6;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 9e-57;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 15
AR194825
LOCUS       AR194825                 447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6350456.
ACCESSION   AR194825
VERSION     AR194825.1  GI:20244262
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
  TITLE     Compositions and methods for the prevention and treatment of M.
            tuberculosis infection
  JOURNAL   Patent: US 6350456-A 4 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             99.6%;  Score 394.4;  DB 6;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 9e-57;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

```
Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 16
AR233097
LOCUS AR233097 447 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 4 from patent US 6458366.
ACCESSION AR233097
VERSION AR233097.1 GI:27275533
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R., Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
TITLE Compounds and methods for diagnosis of tuberculosis
JOURNAL Patent: US 6458366-A 4 01-OCT-2002;
FEATURES Location/Qualifiers
source 1. .447
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 17
AR353302
LOCUS AR353302 447 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 4 from patent US 6592877.
ACCESSION AR353302
VERSION AR353302.1 GI:33759108
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R., Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
TITLE Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL Patent: US 6592877-A 4 15-JUL-2003;
FEATURES Location/Qualifiers
source 1. .447
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 18
AX429596
LOCUS AX429596 447 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 4 from Patent EP1203817.
ACCESSION AX429596
VERSION AX429596.1 GI:21540845
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1
AUTHORS Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L., Vedvick,T.S. and Twardzik,D.R.
TITLE Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL Patent: EP 1203817-A 4 08-MAY-2002;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .447
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 19
AX832581
LOCUS AX832581 447 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 4 from Patent EP1347055.
ACCESSION AX832581
VERSION AX832581.1 GI:39840631
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1
AUTHORS Reed,S.G., Skeiky,Y.A., Campos-Neto,A., Houghton,R.L., Vedvick,T.S., Twardzik,D.R. and Dillon,D.C.
TITLE Compounds for immunotherapy and diagnosis of tuberculosis
JOURNAL Patent: EP 1347055-A 4 24-SEP-2003; CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .447
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 20
BD006325
LOCUS BD006325 447 bp DNA linear PAT 31-JAN-2002
DEFINITION Compounds and methods for diagnosis of Tuberculosis.
ACCESSION BD006325
VERSION BD006325.1 GI:18634696
KEYWORDS JP 2001500383-A/4.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R., Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
TITLE Compounds and methods for diagnosis of Tuberculosis
JOURNAL Patent: JP 2001500383-A 4 16-JAN-2001; CORIXA CORP
COMMENT OS Unidentified
PN JP 2001500383-A/4
PD 16-JAN-2001
PF 07-OCT-1997 JP 1998518432
PR 11-OCT-1996 US 08/729622,13-MAR-1997 US 08/818111 PI STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI ANTONIO CAMPOS NETO,
PI RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI MICHAEL J LODES
PC C12N15/31,C07K14/35,C07K16/12,C12Q1/68,C12N15/62,G01N33/53 CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1. .447
FT /organism='Unidentified'.
FEATURES Location/Qualifiers
source 1. .447
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310
```

```
Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 21
BD006445
LOCUS       BD006445                447 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            Tuberculosis.
ACCESSION   BD006445
VERSION     BD006445.1  GI:18634816
KEYWORDS    JP 2001501832-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001501832-A 4 13-FEB-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001501832-A/4
            PD   13-FEB-2001
            PF   07-OCT-1997 JP 1998518456
            PR   11-OCT-1996 US    08/730510,13-MAR-1997 US    08/818112 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI   ANTONIO CAMPOS
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC   C12N15/31,C07K14/35,A61K39/04,A61K48/00,A61K49/00,C12N15/62,
            PC   C07K19/00,
            PC   G01N33/50,G01N33/60,G01N33/569,C12N1/19,C12N1/20,C12N1/21, PC
             C12N5/10//
            PC   (C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             99.6%;  Score 394.4;  DB 6;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 9e-57;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406


RESULT 22
BD069285
LOCUS       BD069285                447 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            tuberculosis.
ACCESSION   BD069285
VERSION     BD069285.1  GI:22614888
KEYWORDS    JP 2001517069-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.H. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001517069-A 4 02-OCT-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001517069-A/4
            PD   02-OCT-2001
            PF   30-AUG-1996 JP 1997511464
            PR   01-SEP-1995 US    08/523436,22-SEP-1995 US    08/533634 PR
            22-MAR-1996 US    08/620874,05-JUN-1996 US    08/659683 PR
            12-JUL-1996 US    08/680574
            PI   STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
            PI   NETO,
            PI   RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
            C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
            C12N5/10,
            PC   C12N1/21//A61K39/04,(C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Compounds and methods for immunotherapy and diagnosis of  CC
                 tuberculosis
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             99.6%;  Score 394.4;  DB 6;  Length 447;
  Best Local Similarity   99.7%;  Pred. No. 9e-57;
  Matches  395;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190
```

```
Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 23
BD205817
LOCUS BD205817 447 bp DNA linear PAT 17-JUL-2003
DEFINITION Compounds and methods for diagnosis of tuberculosis.
ACCESSION BD205817
VERSION BD205817.1 GI:33015587
KEYWORDS JP 2002530050-A/4.
SOURCE *Mycobacterium tuberculosis*
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 447)
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R., Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
TITLE Compounds and methods for diagnosis of tuberculosis
JOURNAL Patent: JP 2002530050-A 4 17-SEP-2002;
CORIXA CORP
COMMENT OS *Mycobacterium tuberculosis*
PN JP 2002530050-A/4
PD 17-SEP-2002
PF 17-FEB-1999 JP 2000532132
PR 18-FEB-1998 US 09/024753,05-MAY-1998 US 09/072596 PI
STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS PI
NETO,
PI RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK,MICHAEL J
PI LODES,
PI RONALD C HENDRICKSON
PC C12N15/09,C07H21/02,C07H21/04,C07K7/00,C07K14/35,C07K16/12, PC
C07K17/00,
PC C07K19/00,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/04,C12Q1/
PC 68,
PC G01N33/569,G01N33/68//A61K38/00,A61K39/04,A61K39/395,A61K39/
PC 395,A61P31/06,
PC (C12N1/21,C12R1:19),C12N15/00,C12N5/00,A61K37/02 CC
Compounds and methods for diagnosis of tuberculosis. FH Key
Location/Qualifiers
FT source 1. .447
FT /organism='Mycobacterium tuberculosis'.
FEATURES Location/Qualifiers
source 1. .447
/organism="Mycobacterium tuberculosis"
/mol_type="genomic DNA"
/db_xref="taxon:1773"
ORIGIN

Query Match 99.6%; Score 394.4; DB 6; Length 447;
Best Local Similarity 99.7%; Pred. No. 9e-57;
Matches 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

RESULT 24
AR169165
LOCUS AR169165 1872 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 17 from patent US 6290969.
ACCESSION AR169165
VERSION AR169165.1 GI:17906947
KEYWORDS .
SOURCE *Unknown.*
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1872)
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R., Vedvick,T.S. and Twardzik,D.R.
TITLE Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL Patent: US 6290969-A 17 18-SEP-2001;
FEATURES Location/Qualifiers
source 1. .1872
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 98.8%; Score 391.2; DB 6; Length 1872;
Best Local Similarity 99.2%; Pred. No. 2.4e-56;
Matches 393; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117
```

```
Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 25
AR182455
LOCUS       AR182455                1872 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 17 from patent US 6338852.
ACCESSION   AR182455
VERSION     AR182455.1  GI:20225662
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6338852-A 17 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             98.8%;  Score 391.2;  DB 6;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 2.4e-56;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 26
AR194838
LOCUS       AR194838                1872 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 17 from patent US 6350456.
ACCESSION   AR194838
VERSION     AR194838.1  GI:20244275
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
  TITLE     Compositions and methods for the prevention and treatment of M.
            tuberculosis infection
  JOURNAL   Patent: US 6350456-A 17 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             98.8%;  Score 391.2;  DB 6;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 2.4e-56;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


RESULT 27
AR233110
LOCUS       AR233110                1872 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 17 from patent US 6458366.
ACCESSION   AR233110
VERSION     AR233110.1  GI:27275546
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6458366-A 17 01-OCT-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             98.8%;  Score 391.2;  DB 6;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 2.4e-56;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877
```

```
Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

RESULT 28
AR353315

```
LOCUS       AR353315                1872 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 17 from patent US 6592877.
ACCESSION   AR353315
VERSION     AR353315.1  GI:33759121
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6592877-A 17 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            98.8%;  Score 391.2;  DB 6;  Length 1872;
  Best Local Similarity  99.2%;  Pred. No. 2.4e-56;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

RESULT 29
AX429609

```
LOCUS       AX429609                1872 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 17 from Patent EP1203817.
ACCESSION   AX429609
VERSION     AX429609.1  GI:21540858
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: EP 1203817-A 17 08-MAY-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match            98.8%;  Score 391.2;  DB 6;  Length 1872;
  Best Local Similarity  99.2%;  Pred. No. 2.4e-56;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

RESULT 30
AX832594

```
LOCUS       AX832594                1872 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 17 from Patent EP1347055.
ACCESSION   AX832594
VERSION     AX832594.1  GI:39840644
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S., Twardzik,D.R. and Dillon,D.C.
```

```
  TITLE     Compounds for immunotherapy and diagnosis of tuberculosis
  JOURNAL   Patent: EP 1347055-A 17 24-SEP-2003;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             98.8%;  Score 391.2;  DB 6;  Length 1872;
  Best Local Similarity   99.2%;  Pred. No. 2.4e-56;
  Matches  393;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        361 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153


Search completed: April 30, 2004, 07:06:10
Job time : 6636 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: April 30, 2004, 04:41:08 ; Search time 2046 Seconds
(without alignments)
5779.777 Million cell updates/sec

Title: US-09-684-215B-3
Perfect score: 396
Sequence: 1 acggccgcgtccgataactt.....tggccgagggacccccggcc 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	65.2	16.5	289	29	U82114	U82114 U82114 orde
	2	55.6	14.0	603	28	AZ934428	AZ934428 BJ__Ba000
	3	50.4	12.7	925	29	CNS0091P	AL053013 Drosophil
c	4	49.6	12.5	925	29	CNS0091P	AL053013 Drosophil
	5	49.2	12.4	561	28	BH020987	BH020987 P864c.d_L
c	6	48.6	12.3	634	10	BE361689	BE361689 DG1_81_E0
c	7	47.8	12.1	831	28	CC334434	CC334434 OGQAG83TV
	8	47.6	12.0	289	29	P864L	AL390548 Leishmani
	9	47.4	12.0	731	13	BX626840	BX626840 BX626840
	10	47.4	12.0	935	29	CNS006XK	AL066051 Drosophil
c	11	47	11.9	1009	29	CNS010EW	AL098882 Drosophil
	12	46.4	11.7	515	13	BX424977	BX424977 BX424977
	13	45.8	11.6	472	14	CB620050	CB620050 OSIIEa04P
c	14	45.6	11.5	515	13	BX424977	BX424977 BX424977
c	15	45.4	11.5	446	29	CC992789	CC992789 ZUAEQ44TH
c	16	45.4	11.5	459	28	CC349367	CC349367 OGTAB73TV
	17	45.4	11.5	582	29	CG234774	CG234774 OG2BL96TV
	18	45.4	11.5	707	29	CC678226	CC678226 OGWES86TV
c	19	45.4	11.5	777	29	CG326745	CG326745 OG2AH31TV
	20	45.4	11.5	802	29	CG326739	CG326739 OG2AH31TH
c	21	45.4	11.5	830	28	CC331586	CC331586 OGOBH15TV
	22	45.4	11.5	1148	14	CK161968	CK161968 FGAS01455
	23	45.2	11.4	538	10	BE215641	BE215641 HV_CEb000
	24	44.8	11.3	502	9	AU191977	AU191977 AU191977
	25	44.6	11.3	839	29	CNS004NB	AL054280 Drosophil
	26	44.6	11.3	1200	13	BX456467	BX456467 BX456467
c	27	44.6	11.3	1201	13	BX360624	BX360624 BX360624
c	28	44.4	11.2	496	12	BJ307565	BJ307565 BJ307565
c	29	44.4	11.2	504	28	BH190076	BH190076 ATXOC60TF
c	30	44.4	11.2	590	28	BH189840	BH189840 ATXOC41TR
	31	44.4	11.2	632	28	BH189865	BH189865 ATXOD06TF
c	32	44.4	11.2	932	29	CNS0072Q	AL066742 Drosophil
c	33	44.2	11.2	844	29	CNS0052P	AL056652 Drosophil
	34	44	11.1	419	14	CF432426	CF432426 NIT1_13_G
	35	44	11.1	428	14	CF485159	CF485159 POL1_29_F
	36	44	11.1	458	14	CF486070	CF486070 POL1_35_B
	37	44	11.1	493	10	BE355002	BE355002 DG1_10_G0
	38	44	11.1	536	12	BG412839	BG412839 OV2_35_E0
	39	44	11.1	555	14	CF484008	CF484008 POL1_2_B0
	40	44	11.1	575	14	CD222185	CD222185 CCC1_2_H0
	41	44	11.1	587	14	CD229801	CD229801 CCC1_19_A
	42	44	11.1	593	14	CF426951	CF426951 PH1_2_D04
	43	44	11.1	618	14	CD221131	CD221131 CCC1_73_D
	44	44	11.1	644	14	CD226234	CD226234 CCC1_44_E
	45	44	11.1	729	14	CF482692	CF482692 POL1_8_C0

ALIGNMENTS

RESULT 1
U82114
LOCUS U82114 289 bp DNA linear GSS 16-FEB-2001
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone cosmid L-373; contig 64, genomic survey sequence.
ACCESSION U82114
VERSION U82114.1 GI:3647212
KEYWORDS GSS.
SOURCE Mycobacterium leprae
ORGANISM Mycobacterium leprae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
REFERENCE 1 (bases 1 to 289)
AUTHORS Silbaq,F.S., Cho,S.N., Cole,S.T. and Brennan,P.J.
TITLE Characterization of a 34-kilodalton protein of Mycobacterium leprae that is isologous to the immunodominant 34-kilodalton antigen of Mycobacterium paratuberculosis
JOURNAL Infect. Immun. 66 (11), 5576-5579 (1998)
MEDLINE 99003183
PUBMED 9784577
COMMENT Contact: Silbaq FS
Microbiology
Colorado State University
Fort Collins, CO 80523, USA
Eiglmeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use of an ordered cosmid library to deduce the genomic organization of Mycobacterium leprae. Mol. Microbiol. 7 (2), 197-206 (1993)

```
          Class: unknown.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Mycobacterium leprae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1769"
                     /clone="cosmid L-373; contig 64"
                     /clone_lib="ordered cosmid library"
ORIGIN

  Query Match            16.5%;  Score 65.2;  DB 29;  Length 289;
  Best Local Similarity  55.2%;  Pred. No. 0.0046;
  Matches  127;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy        156 CGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAG 215
              ||||||    |  |  |||||   |||      ||| || |   | | | |||| || |
Db         33 CGACAAAGGCACCCCGGGCGCCAAAGTTATGGACGTCGTGGCCGGTGGTGCGGCCGCGAA 92

Qy        216 TCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGC 275
              |  || | |  |||  || | |||  ||||  |||||||| |   | |||||      ||
Db         93 TGCCGCCGTTCCCAAGGGGGTCGTTCTCACTAAGGTCGACGACCGCCTGATCAGTAGCGC 152

Qy        276 CACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTG 335
              |  |||| |||  |  ||  |      |   |||||||||    | ||| ||||||
Db        153 CGACGCGTTGGTCGCTGCCGTGCGGTCCAAGGCACCCGGTGACAAAGTGTCGCTGACCTA 212

Qy        336 GCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTGACATTGGCCGAGG 385
                ||     |||| ||  |||  || || |   | || ||| | | | |||
Db        213 TCAGGATCAGTCTGGTAGCAGTCGCACGGTTCAGGTCACACTCGGCAAGG 262


RESULT 2
AZ934428
LOCUS       AZ934428                603 bp    DNA     linear   GSS 24-APR-2001
DEFINITION  BJ__Ba0002I08r B. japonicum BAC library Bradyrhizobium japonicum
            genomic, genomic survey sequence.
ACCESSION   AZ934428
VERSION     AZ934428.1  GI:13776488
KEYWORDS    GSS.
SOURCE      Bradyrhizobium japonicum
  ORGANISM  Bradyrhizobium japonicum
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1  (bases 1 to 603)
  AUTHORS   Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
            Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
  TITLE     A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
            genome
  JOURNAL   Genome Res. 11 (8), 1434-1440 (2001)
  MEDLINE   21376150
   PUBMED   11483585
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Class: BAC ends
            High quality sequence stop: 553.
FEATURES             Location/Qualifiers
     source          1. .603
                     /organism="Bradyrhizobium japonicum"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:375"
                     /lab_host="E. coli"
                     /clone_lib="B. japonicum BAC library"
                     /note="Vector: pIndigo536; Site_1: HindIII"
ORIGIN

  Query Match            14.0%;  Score 55.6;  DB 28;  Length 603;
  Best Local Similarity  54.4%;  Pred. No. 0.42;

  Matches  112;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        140 TCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGA 199
              |||  |     |  | |  |   || || | | | || || |  ||          |  | ||
Db        363 TCGCCGACAGCCTCGGCATGAAGAAGGCCGAAGGGGCGCTGGTGGCGGAGCCGCAGGCGA 422

Qy        200 GCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCG 259
               || |||||||||| |    |||||||    |||||||||||||||| ||||| ||||||
Db        423 ACGGTCCGGCGGCGAAGGCCGGCATCGAGTCCGGCGACGTGATCACCTCGGTCAACGGCG 482

Qy        260 CTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACG 319
                  |||||    |||  || | | ||   | |  |   |||         ||| |||
Db        483 AATCCGTCAAGGACGCCCGCGAGCTCGCCCGCACCATCGGCGGCATGGCGCCCCGTGCGA 542

Qy        320 TCATCTCGGTGACCTGGCAAACCAAG 345
                ||  |   | | |  |  ||    ||||
Db        543 TCGTGAAGCTTAACGTGCTGCACAAG 568


RESULT 3
CNS0091P
LOCUS       CNS0091P                925 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL053013
VERSION     AL053013.1  GI:4934461
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR19D16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match            12.7%;  Score 50.4;  DB 29;  Length 925;
  Best Local Similarity  15.4%;  Pred. No. 4.8;
  Matches   53;  Conservative  158;  Mismatches  133;  Indels    1;  Gaps    1;

Qy         53 TTCCGATCGGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCA 112
              ||::|  :|:|| :|:|   :|   |:|::::|   :||   : :::  :::: :  |:
Db        553 TTSSGSGYGKGCSSGSGBSCSCCSSCSCSSSSCSCCBCCCCCSCSSYCCSSSBSSSKCSS 612

Qy        113 CCGTTCATATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACG 172
```

```
Db        613 TSBSCSCCCSSKSVCGTSCSSSSSCSSSSSSTSSSSTSSSTSKSSSSGSSSSSSSSYTTS 672

Qy        173 GCGCACGAGTCCAACGCG-TGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACC 231

Db        673 KSTSASGSGSWSAGGGSGSTGSTSSSSSSSSTSTSSSSVSSGSKSSTBSSGSBSSSGSSS 732

Qy        232 GGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGAC 291

Db        733 SSSTSSBBSCTSTSSSSSSSYSSSTCSCCTCCCSYSYSSSTSSSSSTSWGSTSGSSSSSV 792

Qy        292 GCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGC 351

Db        793 GTSSSSDSTSTCCSCCCYMCTCCSTYBMBCYTSTSCGGSSSSSGKGGVTKCGCGGCGSSS 852

Qy        352 GGCACGCGTACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396

Db        853 TNGMBGTSSACSSSSSSCSSSSVSSSSKSSASSSSVSSSGSSGVS 897


RESULT 4
CNS0091P/c
LOCUS       CNS0091P                 925 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL053013
VERSION     AL053013.1  GI:4934461
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .925
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR19D16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match           12.5%;  Score 49.6;  DB 29;  Length 925;
  Best Local Similarity   12.8%;  Pred. No. 7;
  Matches   43;  Conservative  157;  Mismatches  136;  Indels    0;  Gaps    0;

Qy         60 CGGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCA 119

Db        900 SNSSBCSSCSSSBSSSSTSSMSSSSBSSSSGSSSSSSGTSSACVKCNASSSCGCCGCGMA 841

Qy        120 TATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACG 179

Db        840 BCCMCSSSSSCCGSASARGVKVRASGGAGKRGGGSGGASASHSSSSACBSSSSSCSASCW 781

Qy        180 AGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGT 239

Db        780 SASSSSSASSSRSRSGGGAGGSGASSSRSSSSSSSSASAGSVVSSASSSSSSCSSSVSCSS 721

Qy        240 GATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAA 299

Db        720 VASSMSCSSBSSSSASASSSSSSSSASCASCSCCCTSWSCSCSTSASMSAARSSSSSSSS 661

Qy        300 CGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCG 359

Db        660 CSSSSMSASSSASSSSASSSSSSGSSSSSGSACGBSMSSGGGSGSVSASSGMSSSVSSSG 601

Qy        360 TACAGGGAACGTGACATTGGCCGAGGGACCCCCGGC 395

Db        600 GRSSGSGGGGGVGGSGSSSSGSGSSGGSGSVCSCSS 565


RESULT 5
BH020987
LOCUS       BH020987                 561 bp    DNA     linear   GSS 25-MAY-2001
DEFINITION  P864c.d_LEISHPAC-left.1 Leishmania major Friedlin Cosmid Genomic
            Library Leishmania major genomic clone P864c, genomic survey
            sequence.
ACCESSION   BH020987
VERSION     BH020987.1  GI:14202102
KEYWORDS    GSS.
SOURCE      Leishmania major
  ORGANISM  Leishmania major
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Leishmania.
REFERENCE   1  (bases 1 to 561)
  AUTHORS   Myler,P.J., Vogt,C., Cawthra,J., Klacking,M., Marty,A., Mack,J.,
            Munden,H., Nguyen,D., Robertson,L., Sisk,E., Fazelinia,G.,
            Aggarwal,G., Nelson,S., Seyler,A., Worthey,E. and Stuart,K.
  TITLE     Leishmania major Friedlin Cosmid End Sequences
  JOURNAL   Unpublished (2000)
COMMENT     Other_GSSs: P864c.d_LEISHPAC-right.1
            Contact: Myler PJ
            Seattle Biomedical Research Institute
            4 Nickerson Street, Seatttle, WA 98109-1651, USA
            Tel: 206 284-8846
            Fax: 206 284-0313
            Email: mylerpj@sbri.org
            Seq primer: LEISHPAC-left
            Class: PAC end.
FEATURES             Location/Qualifiers
     source          1. .561
                     /organism="Leishmania major"
                     /mol_type="genomic DNA"
                     /strain="Friedlin"
                     /db_xref="taxon:5664"
                     /clone="P864c"
                     /lab_host="E. coli ED8767"
                     /clone_lib="Leishmania major Friedlin Cosmid Genomic
                     Library"
                     /note="Vector: cLHYG; Site_1: BamHI; Genomic DNA from
                     Leishmania major Friedlin was partially digested with
                     Sau3AI, size selected, and ligated with BamHI-digested
                     cLHYG cosmid vector DNA.  9216 clones were picked and
                     arrayed.  Library construction is described in Ivens et
                     al., Genomics Research, 8:135-145 (1998).  The cLHYG
                     vector (Acc. No. CVU59231) is described in Ryan et al,
                     Gene, 131:145-150 (1993)"
ORIGIN

  Query Match           12.4%;  Score 49.2;  DB 28;  Length 561;
  Best Local Similarity   63.6%;  Pred. No. 7.5;
  Matches   75;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;
```

```
Qy        159 CAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCT 218

Db         13 CGAGATCGGCATCGCCGGCGGCCTGGTAGTCGAGGGCGCGTCGGGTCCCGCGGCTCGCGT 72

Qy        219 CGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCC 276

Db         73 CGGCATTCAACCCGGCGACGTGATCCTGGCCTTCAACGGCACGCCGATCAAGTCAGTC 130


RESULT 6
BE361689/c
LOCUS       BE361689               634 bp    mRNA    linear   EST 20-JUL-2000
DEFINITION  DG1_81_E06.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   BE361689
VERSION     BE361689.1  GI:9303246
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
            Pratt,L.H.
  TITLE     An EST database from Sorghum: dark-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: PolyTMix
            High quality sequence start: 67
            High quality sequence stop: 634
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..634
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Dark Grown 1 (DG1)"
                     /note="Organ: 5-day-old dark-grown seedlings; Vector:
                     Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
                     made from poly-A RNA in the cloning vector lambda ZAP II.
                     Clones to be sequenced were prepared by mass excision."
ORIGIN

  Query Match            12.3%;  Score 48.6;  DB 10;  Length 634;
  Best Local Similarity   47.0%;  Pred. No. 10;
  Matches  150;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy         17 ACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATCGGGCAGGCGATGGCGA 76

Db        388 ACGACGTGCCGTCCCAGTAGTGGCAGCGCTTGTGGCCGCCGAGCGCCTGGCCCGTGGCGA 329

Qy         77 TCGCGGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTCATATCGGGCCTACCGCCT 136

Db        328 ACCCCTTCCGGCAGATGGTGCACCGGTGCGGCCCACCGGTAGAAGTCGTCCCTCCCGACG 269

Qy        137 TCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCG 196

Db        268 ACGACGTCGTCGTCGTCGTCGTCTCCTCGGAGTCGGGCGCGGAAAACGCCTCCTGCTGCT 209

Qy        197 GGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACG 256

Db        208 GCGCCGCCGCGGCGGCGGCGTACTGCTCCGGCGTGGGCGGCTTGCGGTGGCTGGACTTGT 149

Qy        257 GCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTG 316

Db        148 GGCCGCCGAGCGCCTGGTACGACGCGAAGGCCTTCCCGCACACCGCGCACCGGAACCGGA 89

Qy        317 ACGTCATCTCGGTGACCTG 335

Db         88 GCTGCTCCTCGTGCGCCAG 70


RESULT 7
CC334434/c
LOCUS       CC334434               831 bp    DNA     linear   GSS 16-MAY-2003
DEFINITION  OGQAG83TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0365N21,
            genomic survey sequence.
ACCESSION   CC334434
VERSION     CC334434.1  GI:30803847
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..831
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0365N21"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match            12.1%;  Score 47.8;  DB 28;  Length 831;
  Best Local Similarity   57.1%;  Pred. No. 15;
  Matches  109;  Conservative    0;  Mismatches   77;  Indels    5;  Gaps    1;

Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223

Db        777 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 723

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283

Db        722 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 663

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCA 343

Db        662 GCGCCTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCCG 603

Qy        344 AGTCGGGCGGC 354

Db        602 CCCCGCGCCGC 592


RESULT 8
P864L
LOCUS       P864L                  289 bp    DNA     linear   GSS 25-JUL-2000
DEFINITION  Leishmania major Friedlin PAC P864 left end-sequence, genomic
```

survey sequence.
ACCESSION AL390548
VERSION AL390548.1 GI:9501524
KEYWORDS GSS.
SOURCE Leishmania major
ORGANISM Leishmania major
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
REFERENCE 1 (bases 1 to 289)
AUTHORS Ivens,A.C., Lewis,S.M., Bagherzadeh,A., Zhang,L., Chan,H.M. and Smith,D.F.
TITLE A physical map of the Leishmania major Friedlin genome
JOURNAL Genome Res. 8 (2), 135-145 (1998)
MEDLINE 98146435
PUBMED 9477341
REFERENCE 2 (bases 1 to 289)
AUTHORS Taylor,R.G., Huckle,E.E.J., Ivens,A.C., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2000) Leishmania major Friedlin genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and alicat@sanger.ac.uk
COMMENT see http://www.ebi.ac.uk/parasites/leish.html
Details of Leishmania sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/L_major/
The primer sequence can be obtained from alicat@sanger.ac.uk.
FEATURES Location/Qualifiers
source 1. .289
/organism="Leishmania major"
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="PAC P864"
ORIGIN

Query Match 12.0%; Score 47.6; DB 29; Length 289;
Best Local Similarity 62.7%; Pred. No. 13;
Matches 74; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
Qy        159 CAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCT 218
              | | | |||||| || ||   |  |   |  || || || |   | ||| ||||    |  |
Db         24 CGAGAGCGGCATCGCCGGCGGCCTGTTAGTCGAGGGCGCGTCGGGTCCCGCGGGTCGCGT 83

Qy        219 CGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCC 276
              ||||||     ||||||||||||||||   ||  || |||||| | ||||||||  || | |
Db         84 CGGCATTCAACCCGGCGACGTGATCCTGGCCTTCAACGGCACGCCGATCAAGTCAGTC 141
```

RESULT 9
BX626840
LOCUS BX626840 731 bp mRNA linear EST 08-AUG-2003
DEFINITION BX626840 NAP1 Anopheles gambiae cDNA clone ANGNP1222C12T7, mRNA sequence.
ACCESSION BX626840
VERSION BX626840.1 GI:33553685
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
REFERENCE 1 (bases 1 to 731)
AUTHORS Lobo,N.L., Gardner,M., Romans,P. and Collins,F.H.
TITLE Anopheles gambiae EST, Center for Tropical Disease Research and Training
JOURNAL Unpublished (2003)
COMMENT Contact: Frank H. Collins
Center for Tropical Disease Research and Training
University of Notre Dame
Notre Dame, IN 46556, USA
Tel: 574-631-9245
Fax: 574-631-3996
Email: frank.h.collins.75@nd.edu.
FEATURES Location/Qualifiers
source 1. .731
/organism="Anopheles gambiae"
/mol_type="mRNA"
/db_xref="taxon:7165"
/clone="ANGNP1222C12T7"
/lab_host="E. coli DH10B"
/clone_lib="NAP1"
/note="Vector: pT7T3D-Pac (Pharmacia); Site_1: NotI; Site_2: EcoRI; ESTs sequenced from the T7 priming site that reads from the 5' end of cDNA. The NAP1 is a directionally cloned and normalized, oligo-T primed cDNA library constructed from a mixture of Anopheles gambiae developmental stages according to: Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."
ORIGIN

Query Match 12.0%; Score 47.4; DB 13; Length 731;
Best Local Similarity 51.2%; Pred. No. 18;
Matches 111; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

```
Qy        156 CGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAG 215
              | ||||||||  |||| | ||             | | |||||     |  ||  |   |
Db         49 CAACAACAACAACAACAGTGCCAACAGTGTCACCCTCGTCGGTCCGTCGGCGATCGAGCG 108

Qy        216 TCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGC 275
              |||  |      | ||  |  |  ||  | |||  |   |  |  |||  ||  ||| | ||
Db        109 GCTCTGTCCAACGACGGCAGCGGTCGCCCCCGTGACGGCCATCGCGACGGCAAACACCGC 168

Qy        276 CACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTG 335
              |||||  ||||||| ||||      ||| |||||  ||||      |  |  ||  | |
Db        169 CACCGGTATGGCGGCCGCGACGGTCGGCCATCACACCGGCATGAGCGGCGGGGAAATCAC 228

Qy        336 GCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 372
              | | |||||   ||  |  | |  ||   | |||  | |
Db        229 GTACACCAACCTGGACTACCCCGGTGTGGTGAAGGAG 265
```

RESULT 10
CNS006XK
LOCUS CNS006XK 935 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC # BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL066051
VERSION AL066051.1 GI:4945019
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 935)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1. .935
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR14N09"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

  Query Match             12.0%;  Score 47.4;  DB 29;  Length 935;
  Best Local Similarity   28.2%;  Pred. No. 19;
  Matches  111;  Conservative   89;  Mismatches  193;  Indels    0;  Gaps    0;

Qy          4 GCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATCGGG 63
              ||: |  :::  : |  |:  | |:|::  : :  :|  |    |||   |:|:  :  :
Db        523 GCSTCCMYMSSSVSCCSCSGTCCGYCSSCTSCKMCSCTYGCKCGCGCGCSTSCSSSSCCS 582

Qy         64 CAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCATATC 123
              : : :     :  :|::| :   | | |     |:| || |:|    |     |      :
Db        583 BBSYSTCSCTBCTKCSSGCSTGSGCTGCCGGGGSGCGGCGSGCGCGCGSGGGCGGGCSSS 642

Qy        124 GGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTC 183
              ||::   : |::  :| ::::  ||| :  |  :        ||:|  ||:||| :| |
Db        643 GGSSGSSSGGSSGSSCGSSSSGSGGGGSGGGCGSGGCGGCGCGSCSGCGSCGCSSGCGSG 702

Qy        184 CAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATC 243
              |     : ||  : | :|::  ||:  ::  : | | :|  : :| :::|:  :
Db        703 CGSGCGSGGGCGSCGCSCSSCGCGSGCSSGSSGCSCCSCCGSCSCSSSSGSCCSCGCGCG 762

Qy        244 ACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGG 303
              ::: :|   : ||||  : |||  :  | :||||  |  |   |:|:  ||||    :||:
Db        763 MSMCSGCGGSSCGGCCSGCCGCSSCGCCSGGCCSGCCGGCGCGSGSCGGCGCGGSSSGGS 822

Qy        304 CATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACA 363
                |  |  |  : |:      ::: : |  |:    ||  | ||:|||:  :| |  :
Db        823 GSCCGSCGGGCGSCCSGGGCGCSSSCSGCGCGSGGSSCCGSGGCGSGCGSGSSGGGSSSS 882

Qy        364 GGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |:|   | |      |:|:| ||: |:|:|| :|
Db        883 GSGSSGGCGSGSGGGSCSGGGGSCCSCSCGCSC 915
```

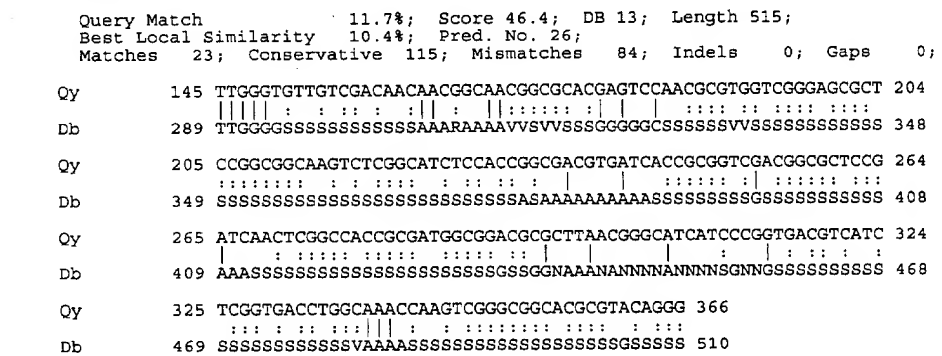
```
RESULT 11
CNS010EW/c
LOCUS       CNS010EW                1009 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN03P19 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL098882
VERSION     AL098882.1  GI:5610493
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1009)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1009
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN03P19"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
ORIGIN

  Query Match             11.9%;  Score 47;  DB 29;  Length 1009;
  Best Local Similarity   30.8%;  Pred. No. 23;
  Matches  111;  Conservative   80;  Mismatches  168;  Indels    1;  Gaps    1;

Qy         37 GGGCAGGGATTCGCCATTCCGATCGGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCG 96
              ::|:|::|:  : :|   : :  ||  :   ||:   ||  | :: : :|:  | |: |
Db       1002 RSGSAVSGMAVVCVCGCGSARCRCGAASCCMCGMACACGCGCMSRARGSAVTCCGGMGCC 943

Qy         97 GGTGGGGGGTCACCCACCGTTCATATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTC 156
              || | :|:  |:|:|   |  :    : ::||  ||:||  :    ::   |  | :| :
Db        942 GGCGMSGSAGCRCSCCGGGGGSCSCCSCSSCCSSCCSCCSSSGGGCSSSGNGMSGSKGGS 883

Qy        157 GACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGT 216
              : :      :|:   |||||  |: : ||   :||  | ||    ::  :||:::|:  :
Db        882 RSMGCGCGGSGSSGGGGCGCSCSCSCCCNCGRCGGCGGCGACGCSSGSSCGSSSGSSGSG 823

Qy        217 CTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCC 276
              : : ||  | |  |  |||  | |    ::|||  ||    :|: ::|  :   ::::|
Db        822 SSSCGCGSCCCGGCGCGCGSGGGGSGGSSSGCGCACGCGCNSGSGSSGGGSGCGGSSSSC 763

Qy        277 ACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGG 336
              ::|:|  :: |::| :||    | :|:  |  :||      |  |  | |::   ||||
Db        762 GSSGSGCSSSGGSRCCSGCGCGGCCSGSTGCGGSCCCCCC-CMCCSCCCCSSCCCCCGGC 704

Qy        337 CAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 396
              |:  |: :| |||| ||   | |    | |  || |     || :| :|: |||| :| |
Db        703 CRSKCSCVGCCGGGGGGGGGGGGGSKGGMGCGCGGGSGSGGGGGSGGSGSBCCCCGSGGC 644
```

```
RESULT 12
BX424977
LOCUS       BX424977                515 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX424977 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA005ZH04
            3-PRIME, mRNA sequence.
ACCESSION   BX424977
VERSION     BX424977.1  GI:30784421
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 515)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 6304.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CL0BA005ZH04FP1&cluster=6304.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CL0BA005ZH04FP1.
FEATURES             Location/Qualifiers
     source          1. .515
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

/db_xref="taxon:9606"
/clone="CL0BA005ZH04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

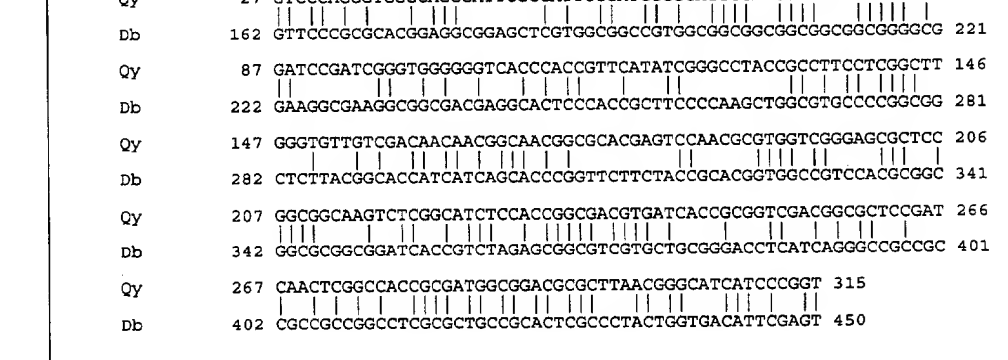
Query Match 11.7%; Score 46.4; DB 13; Length 515;
Best Local Similarity 10.4%; Pred. No. 26;
Matches 23; Conservative 115; Mismatches 84; Indels 0; Gaps 0;



RESULT 13
CB620050
LOCUS CB620050 472 bp mRNA linear EST 08-APR-2003
DEFINITION OSIIEa04P23.f OSIIEa Oryza sativa (indica cultivar-group) cDNA clone OSIIEa04P23 5', mRNA sequence.
ACCESSION CB620050
VERSION CB620050.1 GI:29615037
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 472)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 04 row: P column: 23
Seq primer: gta aaa cga cgg cca gtg.
FEATURES Location/Qualifiers
source 1. .472
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIEa04P23"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIEa"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Lesion Mimic SPL 11"

ORIGIN

Query Match 11.6%; Score 45.8; DB 14; Length 472;
Best Local Similarity 47.4%; Pred. No. 34;
Matches 137; Conservative 0; Mismatches 152; Indels 0; Gaps 0;



RESULT 14
BX424977/c
LOCUS BX424977 515 bp mRNA linear EST 15-MAY-2003
DEFINITION BX424977 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA005ZH04 3-PRIME, mRNA sequence.
ACCESSION BX424977
VERSION BX424977.1 GI:30784421
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 515)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6304.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BA005ZH04FP1&cluster=6304.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CL0BA005ZH04FP1.
FEATURES Location/Qualifiers
source 1. .515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA005ZH04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

```
  Query Match            11.5%;  Score 45.6;  DB 13;  Length 515;
  Best Local Similarity   8.6%;  Pred. No. 37;
  Matches   17;  Conservative  110;  Mismatches   71;  Indels    0;  Gaps    0;

Qy         90 CCGATCGGGTGGGGGGTCACCCACCGTTCATATCGGGCCTACCGCCTTCCTCGGCTTGGG 149
              :::  :: : :::::: : ::: ::: |  |  :::::  :::::  :: :::|   :
Db        512 SSSSSSSCSSSSSSSSSSSSSSSSSSTTTTBSSSSSSSSSSSSSSSSSSSSCNNCSN 453

Qy        150 TGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGC 209
              |          |::|  :::::: :: : ::  ::::|   :::: :::: |::::
Db        452 NNNTNNNNTNTTTNCCSSCSSSSSSSSSSSSSSSSSSSSSSTTTSSSSSSSSSSCSSSS 393

Qy        210 GGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAA 269
              :::   | |     |:|:: :::::: :: :  : :::::: :: :::::: ::: ::
Db        392 SSSSSTTTTTTTTTTSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSBBSS 333

Qy        270 CTCGGCCACCGCGATGGC 287
              : ::||| ||::: ::::
Db        332 SSSSGCCCCCSSSBBSBB 315


RESULT 15
CC992789/c
LOCUS       CC992789                446 bp    DNA     linear   GSS 18-AUG-2003
DEFINITION  ZUAEQ44TH ZM_3.0_4.0_KB Zea mays genomic clone ZMMBPa0044H15,
            genomic survey sequence.
ACCESSION   CC992789
VERSION     CC992789.1  GI:33852665
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: ZUAEQ44TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBPa0044H15"
                     /clone_lib="ZM_3.0_4.0_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 3-4 kb 'unfiltered'
                     genomic DNA library"
ORIGIN

  Query Match            11.5%;  Score 45.4;  DB 29;  Length 446;
  Best Local Similarity   57.5%;  Pred. No. 40;
  Matches  103;  Conservative    0;  Mismatches   71;  Indels    5;  Gaps    1;

Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223
              ||||  |||||||  || ||   ||| |||  |  ||  ||||||    ||  |
Db        306 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 252

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283
              || |||||  |||||||  | |    | || || | | ||| |  || |||||  |||||
Db        251 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 192

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACC 342
              ||   | |      |||||  |  | | | ||||||  ||||      |||  ||   ||
Db        191 GCGCCTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCC 133


RESULT 16
CC349367/c
LOCUS       CC349367                459 bp    DNA     linear   GSS 16-MAY-2003
DEFINITION  OGTAB73TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0387M02,
            genomic survey sequence.
ACCESSION   CC349367
VERSION     CC349367.1  GI:30818774
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 459)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .459
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0387M02"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match            11.5%;  Score 45.4;  DB 28;  Length 459;
  Best Local Similarity   57.5%;  Pred. No. 40;
  Matches  103;  Conservative    0;  Mismatches   71;  Indels    5;  Gaps    1;

Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223
              ||||  |||||||  || ||   ||| |||  |  ||  ||||||    ||  |
Db        446 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 392

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283
              || |||||  |||||||  | |    | || || | | ||| |  || |||||  |||||
Db        391 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 332

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACC 342
              ||   | |      |||||  |  | | | ||||||  ||||      |||  ||   ||
Db        331 GCGCTTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCC 273


RESULT 17
CG234774
LOCUS       CG234774                582 bp    DNA     linear   GSS 22-AUG-2003
DEFINITION  OG2BL96TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0753O24,
            genomic survey sequence.
ACCESSION   CG234774
VERSION     CG234774.1  GI:34134660
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG2BL96TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0753O24"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             11.5%;  Score 45.4;  DB 29;  Length 582;
  Best Local Similarity   57.5%;  Pred. No. 42;
  Matches  103;  Conservative    0;  Mismatches   71;  Indels    5;  Gaps    1;

Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223
              ||||  ||||||||  || ||    ||| |||  |   ||  ||||||    ||   |
Db        280 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 334

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283
              || |||||  |||||||  | |    | || || | | ||| |  || ||||| |||||
Db        335 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 394

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACC 342
                ||   |  |     |||||  |  | | | ||||||  ||||      ||| ||   ||
Db        395 GCGCCTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCC 453


RESULT 18
CC678226
LOCUS       CC678226                 707 bp    DNA     linear   GSS 19-JUN-2003
DEFINITION  OGWES86TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0561P03,
            genomic survey sequence.
ACCESSION   CC678226
VERSION     CC678226.1  GI:32083002
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 707)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .707
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0561P03"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             11.5%;  Score 45.4;  DB 29;  Length 707;
  Best Local Similarity   57.5%;  Pred. No. 44;
  Matches  103;  Conservative    0;  Mismatches   71;  Indels    5;  Gaps    1;

Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223
              ||||  ||||||||  || ||    ||| |||  |   ||  ||||||    ||   |
Db         11 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 65

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283
              || |||||  |||||||  | |    | || || | | ||| |  || ||||| |||||
Db         66 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 125

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACC 342
                ||   |  |     |||||  |  | | | ||||||  ||||      ||| ||   ||
Db        126 GCGCCTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCC 184


RESULT 19
CG326745/c
LOCUS       CG326745                 777 bp    DNA     linear   GSS 26-AUG-2003
DEFINITION  OG2AH31TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0742F14,
            genomic survey sequence.
ACCESSION   CG326745
VERSION     CG326745.1  GI:34244011
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 777)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Other_GSSs: OG2AH31TH
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .777
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0742F14"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

  Query Match             11.5%;  Score 45.4;  DB 29;  Length 777;
  Best Local Similarity   57.5%;  Pred. No. 45;
  Matches  103;  Conservative    0;  Mismatches   71;  Indels    5;  Gaps    1;
```

```
Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223
              ||||  ||||||||  || ||    ||| |||  |   ||  ||||||     ||   |
Db        719 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 665

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283
              || |||||  |||||||  | |     | || || | | ||| |  || ||||| |||||
Db        664 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 605

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACC 342
                ||   |  |   |||||  |  | | | ||||| ||||     ||| ||   ||
Db        604 GCGCCTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCC 546
```

RESULT 20
CG326739
LOCUS CG326739 802 bp DNA linear GSS 26-AUG-2003
DEFINITION OG2AH31TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0742F14, genomic survey sequence.
ACCESSION CG326739
VERSION CG326739.1 GI:34244005
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 802)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OG2AH31TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .802
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0742F14"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 11.5%; Score 45.4; DB 29; Length 802;
Best Local Similarity 57.5%; Pred. No. 45;
Matches 103; Conservative 0; Mismatches 71; Indels 5; Gaps 1;

```
Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223
              ||||  ||||||||  || ||    ||| |||  |   ||  ||||||     ||   |
Db        442 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 496

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283
              || |||||  |||||||  | |     | || || | | ||| |  || ||||| |||||
Db        497 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 556

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACC 342
                ||   |  |   |||||  |  | | | ||||| ||||     ||| ||   ||
Db        557 GCGCCTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCC 615
```

RESULT 21
CC331586/c
LOCUS CC331586 830 bp DNA linear GSS 16-MAY-2003
DEFINITION OGOBH15TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0344C06, genomic survey sequence.
ACCESSION CC331586
VERSION CC331586.1 GI:30800757
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 830)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .830
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMa0344C06"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Query Match 11.5%; Score 45.4; DB 28; Length 830;
Best Local Similarity 57.5%; Pred. No. 46;
Matches 103; Conservative 0; Mismatches 71; Indels 5; Gaps 1;

```
Qy        164 ACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCA 223
              ||||  ||||||||  || ||    ||| |||  |   ||  ||||||     ||   |
Db        389 ACGGGCACGGCGCAGAAGGCCGCTTCGTCGTCCCGGCAGCATCGGCGG-----CTGCCCG 335

Qy        224 TCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGA 283
              || |||||  |||||||  | |     | || || | | ||| |  || ||||| |||||
Db        334 TCGCCACCCTCGACGTGCGCGCGCTCGACGCCGCCCCGCCGCTGCACGCGGCCGCCGCGC 275

Qy        284 TGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACC 342
                ||   |  |   |||||  |  | | | ||||| ||||     ||| ||   ||
Db        274 GCGCCTTCCGGTCCGGCGGGCGCCTCCTCCGCGACGTCCTCTCCTCCTCCTCGCCGTCC 216
```

RESULT 22
CK161968
LOCUS CK161968 1148 bp mRNA linear EST 05-DEC-2003
DEFINITION FGAS014551 Triticum aestivum FGAS: Library 4 Gate 8 Triticum aestivum cDNA, mRNA sequence.
ACCESSION CK161968
VERSION CK161968.1 GI:38990698
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 1148)
AUTHORS Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D., Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A., Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D., Penniket,C., Roach,J.L. and Sarhan,F.
TITLE Functional Genomics of Abiotic Stress In Wheat and Canola Crops
JOURNAL Unpublished (2003)

```
COMMENT     Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy
            (default parameters) has been run on this sequence. Lucy identified
            the region [28,831].
            Plate: L4B003  row: H  column: 15.
FEATURES             Location/Qualifiers
     source          1. .1148
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone_lib="Triticum aestivum FGAS: Library 4 Gate 8"
                     /note="Organ: Crown and leaf; Vector: pCMV.SPORT6; Aerial
                     parts (crown and leaf) of wheat cultivar Norstar during
                     dehydration stress. 8 mRNA populations were combined
                     before constructing the library. The first four come from
                     removing plants from vermiculite (7 day old plants) and
                     incubating them at 20C on the bench without water for 1,
                     2, 3 and 4 days. The last four come from plants grown in
                     soil in a growth chamber after watering is terminated.
                     Four samplings were taken in a two week period; the first
                     after wilting was observed and the last, two weeks later,
                     consisted of live crown and stem tissue (leaf tissue was
                     yellow and dead). First strand synthesis in this library
                     was done in the presence of methylated dCTP thereby
                     protecting from internal cleavage with NotI."
ORIGIN

  Query Match             11.5%;  Score 45.4;  DB 14;  Length 1148;
  Best Local Similarity   47.1%;  Pred. No. 49;
  Matches  139;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy          64 CAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCATATC 123
               | ||| | |  ||||| | || || |||| |||| ||        | | | |    |
Db         255 CTGGCCACGCTGATCGGGCGCGAGCTCCGCGCGGGGGGCTCCGAGCGCCCGGCGCTGCGC 314

Qy         124 GGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTC 183
                ||||  | ||| |  | | |  ||| ||  |  | | |    |
Db         315 TACGGCCACGCCGGCTTCGCCAAGCGCGGGGAGGACTACTTCCTCGTCAAGCCTGACTGC 374

Qy         184 CAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATC 243
               |  |||||   |||   | |  |  | ||       || | || ||     |||||    |
Db         375 CTCCGCGTCCCCGGCGACCCCTCCACCGCCTTCTCCGTCTTCGCCGTGTTCGACGGCCAC 434

Qy         244 ACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGG 303
               | || |||     |  ||   | ||| |  ||     |  | |||  | ||| | |  | | | |
Db         435 AACGGGGTGTCGGCGGCGGTGTTCAGCAAGGAGCACCTGCTGGAGCACGTGATGAGCGCG 494

Qy         304 CATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGC 358
               |  |  || |    || || |   | |  ||||        |  |  | || |||
Db         495 CTCCCGCCGGACATCGGCAGCCGCGAGGACTGGCTGCAGGCGCTGCCCCGCGCGC 549


RESULT 23
BE215641
LOCUS       BE215641                538 bp    mRNA    linear   EST 23-OCT-2001
DEFINITION  HV_CEb0007J02f Hordeum vulgare seedling green leaf EST library
            HVcDNA0005 (Blumeria challenged) Hordeum vulgare subsp. vulgare
            cDNA clone HV_CEb0007J02f, mRNA sequence.
ACCESSION   BE215641
VERSION     BE215641.1  GI:8903169
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D.,
            Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J.,
            Choi,D.W., Fenton,R.D., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Blumeria infected incompatible (Mla6) seedling
            leaf cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 199
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 462.
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="CI16151 (Mla6)"
                     /sub_species="vulgare"
                     /db_xref="taxon:112509"
                     /clone="HV_CEb0007J02f"
                     /tissue_type="seedling green leaf"
                     /lab_host="SOLR"
                     /clone_lib="Hordeum vulgare seedling green leaf EST
                     library HVcDNA0005 (Blumeria challenged)"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
                     C.I. 16151 (Mla6) plants were greenhouse grown in the R
                     Wise lab at Iowa State University, Ames, IA; 7 day old
                     green seedlings were challenged with isolate 5874
                     (AvrMla6) of Blumeria graminis f. sp. hordei, and leaves
                     were harvested 20 and 24 hr post-inoculation and snap
                     frozen; uninoculated leaves were harvested 20 hr
                     post-inoculation (Wei, Wise). In the TJ Close lab at the
                     University of California, Riverside, total RNA was
                     prepared from each sample pool, equal quantities of all
                     three RNA pools were combined, poly(A) RNA was purified
                     from the mixture, one primary unamplified cDNA library was
                     made, and 1 million pfu were in vivo excised to give
                     pBluescript SK(-) cDNA phagemids (Choi, Close). Phagemids
                     were plated and picked at the Clemson University Genomics
                     Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing).
                     Plasmid DNA preparations, DNA sequencing and sequence
                     analysis were performed at CUGI (Wing, Yu, Frisch, Henry,
                     Simmons, Oates, Rambo, Main). The sequence has been
                     trimmed to remove vector sequence and contains a minimum
                     of 100 bases of phred value 20 or above. For more details
                     on library preparation and sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
ORIGIN

  Query Match             11.4%;  Score 45.2;  DB 10;  Length 538;
  Best Local Similarity   48.6%;  Pred. No. 45;
  Matches  122;  Conservative    0;  Mismatches  129;  Indels    0;  Gaps    0;

Qy          54 TCCGATCGGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCAC 113
               | || |     | ||| | |  ||||| | || || ||||  |||| ||       | | |
Db         225 TGCGCTGCCCCTGGCCACGCTGATCGGGCGCGAGCTCCGCGCGGGGGGCTCCGAGCGCCC 284

Qy         114 CGTTCATATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGG 173
                 |  |     |    | ||||  | | | | | | | |  ||| ||     | |
Db         285 GGCGCTGCGCTACGGTCACGCCGGCTTTGTCAAGCGCGGGGAGGACTACTTNCTCGTCAA 344
```

```
Qy        174 CGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGG 233
Db        345 GCCCGACTGCCTTCGCGTACCCGGCGACCCCTCCACCGCCTTCTTCGTCTTCGCCGTGTT 404

Qy        234 CGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGC 293
Db        405 CGACGGGCACAACGGGGTGTCGGGGGCGGGGTTCATCAAGGAGCACCTGCTGGAGCACGT 464

Qy        294 GCTTAACGGGC 304
Db        465 GATGAACGCGC 475
```

```
RESULT 24
AU191977
LOCUS       AU191977                 502 bp    mRNA    linear   EST 14-OCT-2003
DEFINITION  AU191977 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
            cDNA clone PFL024a12_r 5', mRNA sequence.
ACCESSION   AU191977
VERSION     AU191977.1  GI:31930158
KEYWORDS    EST.
SOURCE      Porphyra yezoensis
  ORGANISM  Porphyra yezoensis
            Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
            Porphyra.
REFERENCE   1  (bases 1 to 502)
  AUTHORS   Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
            Tabata,S.
  TITLE     COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
            OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
            FREQUENCY ANALYSIS
  JOURNAL   J. Phycol. 39 (5), 923-930 (2003)
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .502
                     /organism="Porphyra yezoensis"
                     /mol_type="mRNA"
                     /strain="TU-1"
                     /db_xref="taxon:2788"
                     /clone="PFL024a12_r"
                     /dev_stage="sporophytes"
                     /clone_lib="Porphyra yezoensis TU-1 sporophytes"
ORIGIN

  Query Match           11.3%;  Score 44.8;  DB 9;  Length 502;
  Best Local Similarity   47.0%;  Pred. No. 54;
  Matches  139;  Conservative    0;  Mismatches  157;  Indels    0;  Gaps    0;

Qy         57 GATCGGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGT 116
Db         74 GGTGGCGCCGGCGGTGGCGTTCAAGGGGCTGGCCGACTGGGGCAAGTTCTCGTACGCCGC 133

Qy        117 TCATATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGC 176
Db        134 AGGTGACTGGCTCGCCCGCTGCGTGGGCACCGTCACGGCGCCCGCTGGCGGCGCCGGTGC 193

Qy        177 ACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGA 236
Db        194 CGCCGGCCTGCTGGCCGTCAAGGCCGGCAGCGGGTCTGCCCGGGTGGCGGTGGAGGGCAC 253

Qy        237 CGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCT 296
Db        254 CGCCGTGCCCGCGGGCGGTGTGGCGTCGCTCCCCGCCGGCACCCATGCGGTGGAGGTGGT 313

Qy        297 TAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCG 352
Db        314 CTACGCCCACCCCGCCGGCGGCGACGCGGCGCTGTCCGTGACCGCCCCGGCTGGTG 369
```

```
RESULT 25
CNS004NB
LOCUS       CNS004NB                 839 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR10E16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL054280
VERSION     AL054280.1  GI:4931788
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 839)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .839
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR10E16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match           11.3%;  Score 44.6;  DB 29;  Length 839;
  Best Local Similarity   15.9%;  Pred. No. 66;
  Matches   59;  Conservative  150;  Mismatches  162;  Indels    0;  Gaps    0;

Qy          5 CCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATCGGGC 64
Db        465 SMSSSSACSASASSSGCSASSCVAACSSACAGSGASAGSSSSASGSCGAGSSSSAGGSGG 524

Qy         65 AGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCATATCG 124
Db        525 ASGSRSGSGGAGSGGSGGGGSRGSGCGGAVGSASGSSSACSCSASSASSSSSASGCMSCS 584

Qy        125 GGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCC 184
Db        585 SSSSSSSAASCSSSSSSGSVAAVCSGSGSGVGVSAAAAVARSGARGMGGMGAAGSGRAVS 644

Qy        185 AACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCA 244
Db        645 AAASAAASSASSAASVVGAACSSGSGASGSSSAGAARARCASVCASSASSGSCSGCSSSM 704

Qy        245 CCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGC 304
Db        705 CSAVSSGAASSGSASGCCGCSSSSGCSSCSCSSSSMSCSSSCSSSSGCSSVCSCSSCVSS 764

Qy        305 ATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAG 364
Db        765 CSSCSBSSCSCCASSSRASGCSSSSSSSCSSVSCSMSGCMCAMSSASSASSSSAGGSSG 824
```

```
Qy        365 GGAACGTGACA 375
              :::|:: ||::
Db        825 SVVASSCGAVV 835


RESULT 26
BX456467
LOCUS       BX456467                1200 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  BX456467 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP002YE11
            5-PRIME, mRNA sequence.
ACCESSION   BX456467
VERSION     BX456467.1  GI:31032763
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP002AC06QP1.
FEATURES             Location/Qualifiers
     source          1. .1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP002YE11"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             11.3%;  Score 44.6;  DB 13;  Length 1200;
  Best Local Similarity   40.2%;  Pred. No. 71;
  Matches  103;  Conservative   28;  Mismatches  125;  Indels    0;  Gaps    0;

Qy          5 CCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATCGGGC 64
              |:  | || : : |  | :|: : |||   |: |||  |:  :|:|  ||   |   :
Db        918 CSCGGCCCCMCMCCCGCGRGMGKCCCCCCCGBCGGCCCGKCCCVGMCCCCCCCCACMAMM 977

Qy         65 AGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCATATCG 124
             ::|||  |||    :|   ||    ||    |  |  ||||| : ||| |    :    |:
Db        978 MRGCGGGGGCCCCGSCCCCCCCCSCCCCGGCCCGCCGGGGGCSGCCCCCGCGGSGGGGCS 1037

Qy        125 GGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCC 184
             :  | || |   || :||   |||    | | |  |   | |  |||| | |  : ||
Db       1038 SCGCCCCCCCGGSCCSSGGGCNGGGCCCSGGCCCCCCCCCGGCCSCCGGCCCCCCSSCCC 1097

Qy        185 AACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCA 244
               | |  || |:|| | :| |:|||||   | |  |:|  | || | : |  || |  :
Db       1098 CCCCCCGGGCCSGGGGGSCSCSGGCGGGGGGGCCGGSCSCCCCCCCGSCCCCCGGGCCSS 1157

Qy        245 CCGCGGTCGACGGCGC 260
              ||| |  || :  ||:
Db       1158 CCGGGCCCGGSCCCGS 1173


RESULT 27
BX360624/c
LOCUS       BX360624                1201 bp    mRNA    linear   EST 05-MAY-2003
DEFINITION  BX360624 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI072YL05 3-PRIME, mRNA sequence.
ACCESSION   BX360624
VERSION     BX360624.1  GI:30376452
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 4982.r
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI072CF03NP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI072YL05"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             11.3%;  Score 44.6;  DB 13;  Length 1201;
  Best Local Similarity   10.2%;  Pred. No. 71;
  Matches   19;  Conservative  112;  Mismatches   56;  Indels    0;  Gaps    0;

Qy         94 TCGGGTGGGGGGTCACCCACCGTTCATATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTT 153
              ||: :|:  ::::|   : ||:||| | ||::: :  :|:::||   ::::: ::: |
Db       1090 TCSTVTSYCSSBBCSTTSTCCSTTCSTCTCSVSTSSBBCSSBTTTTSSSSSWSSSSSTTS 1031

Qy        154 GTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCA 213
              :|:: :  :  :::: :::::: :: : ::  :::::: ::::: ::::|:::::::::
Db       1030 STBBSBSSSSSSSSBSSSSSSSSSSSSSSSSSSSSSSBBTBSSSSSSSSBTSSSVSSSSV 971

Qy        214 AGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCG 273
              : : :::  : :| ::: :: :: :: : :::::: :: :::::: ::  :     ::
Db        970 SSSSSSSSSSSBCSCSSBBTSSSSSSSWSSSSSSSSSASSSSSSSSSSSSTSKWSSTASS 911

Qy        274 GCCACCG 280
              ::: :::
Db        910 SSSSVSS 904


RESULT 28
BJ307565/c
LOCUS       BJ307565                 496 bp    mRNA    linear   EST 09-APR-2002
DEFINITION  BJ307565 Y. Ogihara unpublished cDNA library, Wh_yd Triticum
            aestivum cDNA clone whyd11p23 3', mRNA sequence.
ACCESSION   BJ307565
VERSION     BJ307565.1  GI:20116389
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   Ogihara,Y. and Murai,K.
  TITLE     Expressed genes in Triticum aestivum
```

```
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .496
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="whyd11p23"
                     /tissue_type="spikelet at late flowering"
                     /dev_stage="Feekes' scale 6"
                     /clone_lib="Y. Ogihara unpublished cDNA library, Wh_yd"
ORIGIN

  Query Match             11.2%;  Score 44.4;  DB 12;  Length 496;
  Best Local Similarity   49.2%;  Pred. No. 64;
  Matches  117;  Conservative    0;  Mismatches  121;  Indels    0;  Gaps    0;

Qy        148 GGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCG 207
              | ||  |  ||  || | |  | ||  || ||| |  | |    || |  ||| | |
Db        457 GCTGGAGGTGATCACCAACGGCAAGGTACAAGAGGCGTGCTCCACCGCGTGGTCGCGCAG 398

Qy        208 GCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATC 267
              || |   |   | | || |||| || |  | |   || | ||| |  || |||     |||
Db        397 GCCGACGGCAACCGGATGTCCATCGCCTCCTTCTACAACCCGGCCAGCGACGCGGTCATC 338

Qy        268 AACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCG 327
                | ||||  | ||| |||||||| |||      |||      | ||| |   |||  ||  |
Db        337 TTCCCGGCGCCGGCGCTGGCGGAGGCGGCCGGCGGTGCGTACCCCAGGTTCGTGTTCGAG 278

Qy        328 GTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTGACATTGGCCGAGG 385
              |    |  |  |      |  | ||  || |      | || | | |   | |||||
Db        277 GACTACATGAAGCTGTACGTGCGCCACAAGTTCGATGACAAGGAGCCCAGGTTCGAGG 220


RESULT 29
BH190076/c
LOCUS       BH190076                504 bp    DNA     linear   GSS 19-OCT-2001
DEFINITION  ATXOC60TF ATXO Arabidopsis thaliana genomic clone ATXOC60, genomic
            survey sequence.
ACCESSION   BH190076
VERSION     BH190076.1  GI:16302878
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 504)
  AUTHORS   Kaul,S., Town,C.D., Bowman,C.L., Van Aken,S.E., Utterback,T.V.,
            Feldblyum,T.V. and Fraser,C.M.
  TITLE     Survey sequencing of Arabidopsis thaliana BAC T24G4
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: ATXOC60TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            From Wash. U contig 1142. Caution: the DNA in this BAC may be from
            some non-Arabidopsis source
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="ATXOC60"
                     /clone_lib="ATXO"
                     /note="Vector: pHOS2; Site_1: BstXI; 2-3 kb sheared BAC
                     DNA inserted into pHOS2 using BstXI linkers"
ORIGIN

  Query Match             11.2%;  Score 44.4;  DB 28;  Length 504;
  Best Local Similarity   53.4%;  Pred. No. 64;
  Matches   93;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy        155 TCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAA 214
              ||| |  | ||| || ||||          ||  | ||  | |  |  ||| |  |  ||
Db        401 TCGGCTTCGACGTCATCGGCATGACCAACCTCCCCGAAGCCAAGCTCGCCCGCGAAGCCG 342

Qy        215 GTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGG 274
                ||| | || ||||  || | ||||||||||     |||| ||       | | ||   |
Db        341 AGATCGCCCTCGCCACCCTCGCCATGATCACCGACTACGACTGCTGGAAAACCGACGAAG 282

Qy        275 CCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGG 328
              || |||  |  || || ||  | | ||  || |  | ||||  |  ||
Db        281 CCCCCGTCACCGCCGAAGCCGTCATCGCCCACCTCCACGCCAATGTCAGCATGG 228


RESULT 30
BH189840/c
LOCUS       BH189840                590 bp    DNA     linear   GSS 19-OCT-2001
DEFINITION  ATXOC41TR ATXO Arabidopsis thaliana genomic clone ATXOC41, genomic
            survey sequence.
ACCESSION   BH189840
VERSION     BH189840.1  GI:16302367
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Kaul,S., Town,C.D., Bowman,C.L., Van Aken,S.E., Utterback,T.V.,
            Feldblyum,T.V. and Fraser,C.M.
  TITLE     Survey sequencing of Arabidopsis thaliana BAC T24G4
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: ATXOC41TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            From Wash. U contig 1142. Caution: the DNA in this BAC may be from
            some non-Arabidopsis source
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="ATXOC41"
                     /clone_lib="ATXO"
                     /note="Vector: pHOS2; Site_1: BstXI; 2-3 kb sheared BAC
                     DNA inserted into pHOS2 using BstXI linkers"
ORIGIN

  Query Match             11.2%;  Score 44.4;  DB 28;  Length 590;
  Best Local Similarity   53.4%;  Pred. No. 66;
  Matches   93;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy        155 TCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAA 214
              ||| |  | ||| || ||||          ||  | ||  | |  |  ||| |  |  ||
```

```
Db        216 TCGGCTTCGACGTCATCGGCATGACCAACCTCCCCGAAGCCAAGCTCGCCCGCGAAGCCG 157

Qy        215 GTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTCGACGGCGCTCCGATCAACTCGG 274
              ||| | || ||||| || | ||||||||||     |||| ||       | | ||   |
Db        156 AGATCGCCCTCGCCACCCTCGCCATGATCACCGACTACGACTGCTGGAAAACCGACGAAG 97

Qy        275 CCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGG 328
              || |||  |  || || ||  | | ||  || |  | ||   | |||| |  ||
Db         96 CCCCCGTCACCGCCGAAGCCGTCATCGCCCACCTCCACGAAAATGTCAGCATGG 43
```

Search completed: April 30, 2004, 07:40:32
Job time : 2052 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:30:12 ; Search time 364.096 Seconds
(without alignments)
1493.479 Million cell updates/sec

Title: US-09-684-215B-18
Perfect score: 653
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13179/app_query.fasta_1.1180
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_819_@runat_29042004_061304_13179 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	653	100.0	675	4	AAH93896	Aah93896 Ra12-P510
2	653	100.0	675	4	AAS64132	Aas64132 Human /M.
3	653	100.0	675	5	ACA59940	Aca59940 Prostate
4	653	100.0	675	6	ABL95503	Abl95503 Ra12- P51
5	653	100.0	675	7	ACC95667	Acc95667 Prostate
6	653	100.0	675	9	ADB14272	Adb14272 Human pro
7	653	100.0	822	6	ABK39769	Abk39769 DNA encod
8	653	100.0	822	7	ACA12098	Aca12098 Human lun
9	653	100.0	822	7	ACA03284	Aca03284 Lung canc
10	653	100.0	894	6	ABK27798	Abk27798 Human col
11	653	100.0	900	6	ABL49257	Abl49257 Ra12/C-te
12	653	100.0	900	6	ABQ92443	Abq92443 Human lun
13	653	100.0	900	8	ADA28443	Ada28443 Lung tumo
14	653	100.0	915	4	AAH93905	Aah93905 Ra12-P775
15	653	100.0	915	4	AAS64141	Aas64141 Human /M.
16	653	100.0	915	5	ACA59949	Aca59949 Prostate
17	653	100.0	915	6	ABL95512	Abl95512 Ra12- P77
18	653	100.0	915	7	ACC95676	Acc95676 Prostate
19	653	100.0	915	9	ADB14284	Adb14284 Human pro
20	653	100.0	945	6	ABK39768	Abk39768 DNA encod
21	653	100.0	945	7	ACA12097	Aca12097 Human lun
22	653	100.0	945	7	ACA03283	Aca03283 Lung canc
23	653	100.0	1012	6	ABL49256	Abl49256 Ra12/N-te
24	653	100.0	1012	6	ABQ92442	Abq92442 Human lun
25	653	100.0	1012	8	ADA28441	Ada28441 Lung tumo
26	653	100.0	1035	6	ABK69714	Abk69714 DNA encod
27	653	100.0	1035	7	ABZ33699	Abz33699 Human col
28	653	100.0	1035	9	ADB67590	Adb67590 Human RA1
29	653	100.0	1155	6	ABK39775	Abk39775 DNA encod
30	653	100.0	1155	7	ACA12104	Aca12104 Human Ra1
31	653	100.0	1155	7	ACA03290	Aca03290 Lung canc
32	653	100.0	1203	4	AAH93917	Aah93917 Ra12-P501
33	653	100.0	1203	4	AAS64153	Aas64153 Human /M.
34	653	100.0	1203	5	ACA59961	Aca59961 Prostate
35	653	100.0	1203	6	ABL95524	Abl95524 Ra12-P501
36	653	100.0	1203	7	ACC95688	Acc95688 Prostate
37	653	100.0	1203	9	ADB14301	Adb14301 Human pro
38	653	100.0	1263	6	ABK69715	Abk69715 DNA encod
39	653	100.0	1263	9	ADB67591	Adb67591 Human RA1
40	653	100.0	1464	4	AAH56353	Aah56353 DNA encod
41	653	100.0	1464	6	ABL92582	Abl92582 Chlamydia
42	653	100.0	1557	4	AAH56341	Aah56341 DNA encod
43	653	100.0	1557	6	ABL92570	Abl92570 Chlamydia
44	653	100.0	1578	3	AAA64764	Aaa64764 C. pneumo
45	653	100.0	1578	4	AAH56267	Aah56267 Chlamydia

ALIGNMENTS

RESULT 1
AAH93896
ID AAH93896 standard; cDNA; 675 BP.
XX
AC AAH93896;
XX
DT 04-OCT-2001 (first entry)
XX
DE Ra12-P510S-C construct cDNA sequence.
XX
KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytostatic; gene therapy; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US001574.
XX
PR 14-JAN-2000; 2000US-00483672.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX
DR WPI; 2001-425873/45.
XX
PT New polynucleotide encoding a prostate-specific protein, for diagnosing,

PT monitoring and treating prostate cancer in a patient and for use in
PT vaccines.
XX
PS Claim 8; Page 492-493; 543pp; English.
XX
CC The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated T cells
CC prepared using (I) or (II) are used treat cancer in a patient. (I) and
CC the antibodies are also used in the detection of cancer in a patient. The
CC cancer that is diagnosed or treated is particularly prostate cancer. (I)
CC and (II) can be used in vaccines. The antibodies or (I) can be used for
CC monitoring the progression of cancer in a patient. (I) and (II) can also
CC be used to improve diagnostic and therapeutic methods for prostate
CC cancer. They can indicate the level of metastasis as well as the prostate
CC volume. AAH93357 to AAH93944 and AAM01115 to AAM01318 represent
CC polynucleotide and amino acid sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.71e-58	Length:	675
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x AAH93896 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 2
AAS64132
ID AAS64132 standard; cDNA; 675 BP.
XX
AC AAS64132;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human /M. tuberculosis Ra12 fusion protein RA12-P510S-C cDNA.
XX
KW Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS Homo sapiens.
OS Microbacterium; tuberculosis.
OS Synthetic.
OS Chimeric.
XX
PN WO200173032-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US009919.
XX
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR WPI; 2001-639232/73.
DR P-PSDB; AAU69899.
XX
PT New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer.
XX
PS Example 17; Page 532; 579pp; English.
XX
CC The invention relates to isolated prostate-specific polynucleotides,
CC polypeptides, fusion proteins of the polypeptides, antibodies raised
CC against the polypeptides (or antigenic epitopes derived from them) and
CC antigen-presenting cells expressing the polypeptides. The antibodies are
CC useful for detecting the presence of cancer, especially prostate cancer.
CC The polypeptides, polynucleotides and the antigen-presenting cells are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein, and for inhibiting the development of cancer especially prostate
CC cancer. Compositions comprising the polynucleotide and/or polypeptide are
CC useful for stimulating an immune response, and for treating cancer. The
CC oligonucleotide is useful for detecting cancer. The present sequence is a
CC prostate specific polynucleotide of the invention
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.71e-58	Length:	675
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x AAS64132 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201
```

```
Qy          61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy          81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy         101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy         121 ThrLeuAlaGluGlyProProAla 128
               ||||||||||||||||||||||||
Db         382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 3
ACA59940
ID ACA59940 standard; cDNA; 675 BP.
XX
AC ACA59940;
XX
DT 10-JUN-2003 (first entry)
XX
DE Prostate cancer therapy associated cDNA #647.
XX
KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;
KW immunogen; cancer; prostate specific antigen; PSA;
KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;
KW PSMA; gene; ss.
XX
OS Homo sapiens.
XX
PN US2002192763-A1.
XX
PD 19-DEC-2002.
XX
PF 29-JUN-2001; 2001US-00895793.
XX
PR 04-OCT-1999; 99US-0157455P.
PR 04-OCT-2000; 2000US-00679272.
PR 28-MAR-2001; 2001US-00822827.
XX
PA (XUJJ/) XU J.
PA (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (HARL/) HARLOCKER S L.
PA (JIAN/) JIANG Y.
PA (KALO/) KALOS M D.
PA (FANG/) FANGER G R.
PA (RETT/) RETTER M W.
PA (STOL/) STOLK J A.
PA (DAYC/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
PA (HEND/) HENDERSON R A.
PA (HURA/) HURAL J.
PA (MCNE/) MCNEILL P D.
PA (HOUG/) HOUGHTON R L.
PA (DBAS/) Y DE BASSOLS C V.
PA (FOYT/) FOY T M.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;
PI Mcneill PD, Houghton RL, Y De BassolsCV, Foy TM;
XX
DR WPI; 2001-245062/25.
XX
PT Prostate specific protein and its encoding polynucleotide, useful for the
PT treatment and diagnosis of prostate cancer.
XX
PS Example 17; SEQ ID NO 822; 85pp; English.
XX
CC The invention describes a fusion protein comprising at least one amino
CC acid sequence of immunogenic portions of any of the 3 sequences not
CC defined in the specification, or sequences having at least 70 or 90 %
CC sequence identity to any one of the 35 sequences defined in the USPTO web
CC site, which is encoded by any of the 4 nucleotide sequences not defined
CC in the specification. The fusion protein, composition and methods are
CC useful for diagnosing, preventing and/or treating cancer, particularly
CC prostate cancer. The proteins are useful as markers to indicate the
CC presence or absence of cancer. This sequence represents a prostate cancer
CC therapy associated cDNA. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=US20020192763
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.71e-58	Length:	675
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-684-215B-18 (1-128) x ACA59940 (1-675)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy          41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy          61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy          81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy         101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy         121 ThrLeuAlaGluGlyProProAla 128
               ||||||||||||||||||||||||
Db         382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 4
ABL95503
ID ABL95503 standard; cDNA; 675 BP.
XX
AC ABL95503;
XX
DT 29-AUG-2003 (revised)
DT 19-JUL-2002 (first entry)
XX
DE Ra12- P510S-C construct cDNA sequence SEQ ID NO 822.
XX
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW gene therapy; gene; ss.
XX

```
OS   Mycobacterium tuberculosis.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-00759143.
XX
PR   25-FEB-1997;   97US-00806099.
PR   01-AUG-1997;   97US-00904804.
PR   10-FEB-1998;   98US-00020956.
PR   25-FEB-1998;   98US-00030607.
PR   14-JUL-1998;   98US-00115453.
PR   23-SEP-1998;   98US-00159812.
PR   15-JAN-1999;   99US-00232149.
PR   09-APR-1999;   99US-00288946.
PR   13-JUL-1999;   99US-00352616.
PR   12-NOV-1999;   99US-00439313.
PR   18-NOV-1999;   99US-00443686.
PR   14-JAN-2000; 2000US-00483672.
PR   27-MAR-2000; 2000US-00536857.
PR   09-MAY-2000; 2000US-00568100.
PR   12-MAY-2000; 2000US-00570737.
PR   13-JUN-2000; 2000US-00593793.
PR   27-JUN-2000; 2000US-00605783.
PR   10-AUG-2000; 2000US-00636215.
PR   29-AUG-2000; 2000US-00651236.
PR   06-SEP-2000; 2000US-00657279.
PR   02-OCT-2000; 2000US-00679426.
PR   10-OCT-2000; 2000US-00685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI   Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI   Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer.
XX
PS   Example 17; SEQ ID NO 822; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention. (Updated on 29-AUG-2003 to standardise OS
CC   field)
XX
SQ   Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.71e-58        Length:         675
```

```
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x ABL95503 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

```
RESULT 5
ACC95667
ID   ACC95667 standard; cDNA; 675 BP.
XX
AC   ACC95667;
XX
DT   28-AUG-2003  (first entry)
XX
DE   Prostate tumour specific cDNA sequence SEQ ID 822.
XX
KW   Cytostatic; gene therapy; prostate-specific protein; PSP; human;
KW   immune response; prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200289747-A2.
XX
PD   14-NOV-2002.
XX
PF   09-MAY-2002; 2002WO-US014753.
XX
PR   09-MAY-2001; 2001US-00852911.
PR   29-JUN-2001; 2001US-00895814.
PR   10-DEC-2001; 2001US-00012896.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA;
PI   Kalos MD, Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS;
PI   Carter D, Li SX, Wang A, Skeiky YAW, Hepler WT, Hural J;
PI   Mcneill PD, Houghton RL, Vinals Y De BassolsC, Foy TM, Watanabe Y;
PI   Deng T;
XX
DR   WPI; 2003-167130/16.
XX
PT   New prostate-specific proteins and genes, useful in gene therapy,
PT   particularly for stimulating an immune response in a patient, or treating
```

PT prostate cancer in a patient, as well as for diagnosing prostate cancer
PT in a patient.
XX
PS Example 17; Page 601-602; 691pp; English.
XX
CC The present invention relates to novel prostate-specific proteins (PSP)
CC and their coding sequences. The PSPs and their coding sequences are
CC useful for stimulating an immune response in a patient, or for treating
CC prostate cancer in a patient and for determining, detecting or diagnosing
CC the presence of a cancer in a patient. The present sequence was used to
CC illustrate the invention
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.71e-58	Length:	675
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-09-684-215B-18 (1-128) x ACC95667 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 6
ADB14272
ID ADB14272 standard; cDNA; 675 BP.
XX
AC ADB14272;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human prostate specific protein P510S-RA12 construct C cDNA.
XX
KW Human; ss; prostate specific cDNA; cytostatic; immunostimulant;
KW gene therapy; cell therapy; vaccine; T-cell epitope;
KW class I major histocompatibility complex allele; MHC; prostate cancer;
KW tumour; antigen presenting cell; gene; fusion protein.
XX
OS Chimeric.
OS Homo sapiens.
OS Mycobacterium tuberculosis.
XX
PN US2003185830-A1.
XX
PD 02-OCT-2003.
XX
PF 12-NOV-2002; 2002US-00294025.
XX
PR 25-FEB-1997; 97US-00806099.
PR 01-AUG-1997; 97US-00904804.
PR 09-FEB-1998; 98US-00020956.
PR 25-FEB-1998; 98US-00030607.
PR 14-JUL-1998; 98US-00115453.
PR 23-SEP-1998; 98US-00159812.
PR 15-JAN-1999; 99US-00232149.
PR 09-APR-1999; 99US-00288946.
PR 13-JUL-1999; 99US-00352616.
PR 12-NOV-1999; 99US-00439313.
PR 18-NOV-1999; 99US-00443686.
PR 14-JAN-2000; 2000US-00483672.
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
PR 12-JAN-2001; 2001US-00759143.
PR 09-FEB-2001; 2001US-00780669.
PR 09-MAY-2001; 2001US-00852911.
PR 29-JUN-2001; 2001US-00895814.
PR 10-DEC-2001; 2001US-00012896.
PR 09-MAY-2002; 2002US-00144678.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Stolk JA, Kalos MD;
XX
DR WPI; 2003-756193/71.
DR P-PSDB; ADB14275.
XX
PT New isolated polypeptide for use in a vaccine for stimulating an immune
PT response, or for treating or diagnosis cancer, preferably prostate
PT cancer.
XX
PS Example 17; Page; 101pp; English.
XX
CC The invention relates to an isolated polypeptide comprising no more than
CC 11-542 amino acids of ADB13563 comprising a sequence ADB14487. The
CC peptides comprise a fragment ADB13563 of that contain naturally processed
CC T-cell epitopes for 3 class I major histocompatibility complex (MHC)
CC alleles. ADB13563 is a polypeptide encoded by a human prostate specific
CC cDNA, one of 648 disclosed as new. Also included are nucleic acids
CC encoding the proteins and peptides, expression vectors, a host cell
CC transformed with the vector, an isolated antibody (or antigen binding
CC fragment) that specifically binds to the protein or peptide, detecting
CC the presence of a cancer in a patient (comprising contacting a patient
CC sample with a binding agent that binds to the peptides or a polypeptide
CC appearing as ADB13558, detecting the amount of polypeptide that binds to
CC the agent and comparing the amount of polypeptide to a predetermined cut-
CC off value to determine the presence of cancer), a fusion protein
CC comprising the peptides or proteins, stimulating or expanding T cells
CC specific for a tumour protein comprising contacting T cells with the
CC peptides or the isolated T cell population, treating prostate cancer in a
CC patient comprising administering a composition comprising the peptides,
CC nucleic acids, antibodies or compounds, determining the presence of a
CC cancer in a patient and treating prostate cancer in a patient comprising
CC incubating cluster of differentiation (CD4+) and/or CD8+ T cells isolated
CC from a patient with the peptides or antigen presenting cells that express
CC (the peptides so that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The peptides (or an oligonucleotide
CC that hybridises to nucleic acid encoding them), is used to detect the
CC presence of cancer in a patient. The peptides, nucleic acids encoding, or

```
CC   antigen-presenting cells expressing the nucleic acid, are used to
CC   stimulate or expand T cells specific for a tumour protein. The peptides,
CC   nucleic acids, antibodies, fusion proteins, T cell populations or antigen
CC   presenting cells are used to stimulate an immune response or treat
CC   prostate cancer in a patient. The present sequence is a cDNA encoding a
CC   fusion protein comprising a prostate specific protein. Note: Except where
CC   otherwise indicated, the sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030185830.
XX
SQ   Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.71e-58        Length:         675
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       9               Gaps:           0

US-09-684-215B-18 (1-128) x ADB14272 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 7
ABK39769
ID   ABK39769 standard; cDNA; 822 BP.
XX
AC   ABK39769;
XX
DT   21-MAY-2002  (first entry)
XX
DE   DNA encoding lung tumour protein P801P ORF5 and Ra12 fusion protein.
XX
KW   Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW   gene; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200204514-A2.
XX
PD   17-JAN-2002.
XX
PF   10-JUL-2001; 2001WO-US022058.
XX
PR   11-JUL-2000; 2000US-00614124.
PR   29-AUG-2000; 2000US-00651563.
PR   08-SEP-2000; 2000US-00658824.
PR   26-SEP-2000; 2000US-00671325.
PR   06-OCT-2000; 2000US-00677419.
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Watanabe Y,  Henderson RA,  Johnson JC,  Retter MW;
PI   Marnerakis M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
PI   Wang A,  Fanger N,  Switzer A,  Mcneill PD,  Clapper JD;
XX
DR   WPI; 2002-164634/21.
DR   P-PSDB; AAU85587.
XX
PT   Novel polynucleotide encoding a lung tumor polypeptide useful for
PT   stimulating and/or expanding T cells specific for a tumor protein.
XX
PS   Example 8; SEQ ID NO 1862; 223pp; English.
XX
CC   The invention describes an isolated polynucleotide and polypeptide useful
CC   for stimulating and/or expanding T cells specific for a tumour protein
CC   for determining the presence of a cancer in a patient. A composition
CC   containing the polynucleotide and/or polypeptide is useful for treating a
CC   lung cancer in a patient. The polypeptide is useful for removing tumour
CC   cells from a biological sample. The polynucleotide is also useful as
CC   probe or primer to detect the level of mRNA encoding a tumour protein.
CC   This sequence encodes a lung tumour associated protein or protein
CC   fragment, described in the method of the invention. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.44e-58        Length:         822
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-18 (1-128) x ABK39769 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
```

```
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 8
ACA12098
ID   ACA12098 standard; cDNA; 822 BP.
XX
AC   ACA12098;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Human lung cancer protein L801P ORF5/Ra12 fusion protein cDNA.
XX
KW   Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW   T cell expansion; CD4; CD8; RA12; gene.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US2002197669-A1.
XX
PD   26-DEC-2002.
XX
PF   03-MAY-2001; 2001US-00849626.
XX
PR   13-DEC-2000; 2000US-00736457.
XX
PA   (BANG/) BANGUR C S.
PA   (FANG/) FANGER G R.
PA   (WANG/) WANG A.
PA   (WANG/) WANG T.
PA   (SWIT/) SWITZER A P.
PA   (MCNE/) MCNEILL P D.
PA   (CLAP/) CLAPPER J D.
XX
PI   Bangur CS,  Fanger GR,  Wang A,  Wang T,  Switzer AP,  Mcneill PD;
PI   Clapper JD;
XX
DR   WPI; 2003-352750/33.
DR   P-PSDB; ABU69562.
XX
PT   Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT   detecting the presence of lung cancer in a patient, and in pharmaceutical
PT   compositions, e.g. vaccines, for treating lung cancer.
XX
PS   Example 8; Page; 72pp; English.
XX
CC   The invention relates to a polynucleotide encoding a lung tumour protein,
CC   comprising a sequence selected from any of the 14 sequences mentioned in
CC   the specification, or a sequence (S2) mentioned in specification,
CC   complement of S1, sequences consisting of at least 20 contiguous residues
CC   of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC   90%, identity to S1, or degenerate variants of S1. Also included are an
CC   isolated polypeptide (comprising a sequence (S3) selected from any one of
CC   the 4 amino acid sequences mentioned in the specification, a sequence
CC   encoded by the polynucleotide, or sequences having at least 70%,
CC   preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC   expression vector comprising the polynucleotide operably linked to an
CC   expression control sequence, a host cell transformed or transfected with
CC   the vector, an isolated antibody (or its antigen-binding fragment) that
CC   specifically binds to the polypeptide, detecting the presence of a cancer
CC   in a patient, a fusion protein comprising the polypeptide, an
CC   oligonucleotide that hybridises to S1 under moderately stringent
CC   conditions, stimulating and/or expanding T cells specific for a tumour
CC   protein (comprising contacting T cells with the polynucleotide, protein
CC   or antigen-presenting cells, under conditions and for a time sufficient
CC   to permit the stimulation and/or expansion of T cells) and inhibiting the
CC   development of a cancer in a patient (by incubating CD4^+ and/or CD8^+ T
CC   cells isolated from a patient with the polynucleotide, protein or antigen
CC   presenting cells that express the polynucleotide, such that T cells
CC   proliferate, administering to the patient an effective amount of the
CC   proliferated T cells, and thus inhibiting the development of a cancer in
CC   the patient. The polynucleotide, protein ands cells are useful in a
CC   composition for stimulating an immune response in a patient, and for
CC   treating a cancer in a patient (particularly lung cancer). The
CC   oligonucleotide is useful for determining the presence of a cancer in a
CC   patient. The protein and oligonucleotides are useful in pharmaceutical
CC   compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC   or primer for nucleic acid hybridisation, and in the design and
CC   preparation of ribozyme molecules for inhibiting expression of tumour
CC   polypeptides and proteins in tumour cells. An amplified portion of the
CC   polynucleotide is useful for isolating a full-length gene from a suitable
CC   library. The present sequence encodes a fusion protein of human RA12 with
CC   the protein product of a cDNA (full length, extended or partial) isolated
CC   from a library derived from lung tumour/cancer cells. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from the USPTO at
CC   seqdata.uspto.gov/sequence.html?DocId=20020197669
XX
SQ   Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.44e-58        Length:          822
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       7               Gaps:            0

US-09-684-215B-18 (1-128) x ACA12098 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 9
ACA03284
ID   ACA03284 standard; DNA; 822 BP.
XX
AC   ACA03284;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Lung cancer therapyand diagnosis associated DNA #5.
XX
KW   Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   US2002172952-A1.
XX
PD   21-NOV-2002.
XX
```

```
PF   10-JUL-2001; 2001US-00902941.
XX
PR   30-JUN-1999;   99US-00346492.
PR   15-OCT-1999;   99US-00419356.
PR   17-DEC-1999;   99US-00466867.
PR   30-DEC-1999;   99US-00476300.
PR   06-MAR-2000; 2000US-00519642.
PR   22-MAR-2000; 2000US-00533077.
PR   10-APR-2000; 2000US-00546259.
PR   27-APR-2000; 2000US-00560406.
PR   05-JUN-2000; 2000US-00589184.
PR   11-JUL-2000; 2000US-00614124.
PR   29-AUG-2000; 2000US-00651563.
PR   08-SEP-2000; 2000US-00658824.
PR   26-SEP-2000; 2000US-00671325.
PR   06-OCT-2000; 2000US-00677419.
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Henderson RA,  Wang T,  Watanabe Y,  Johnson JC,  Retter MW;
PI   Durham M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
XX
DR   WPI; 2003-328427/31.
XX
PT   New polynucleotide, useful for preparing a composition for treating or
PT   inhibiting development of cancer, e.g. lung cancer.
XX
PS   Example 8; SEQID NO 1862; 82pp; English.
XX
CC   The invention describes an isolated polynucleotide comprising one of 32
CC   sequences, complement or degenerate variants of them. The polynucleotide
CC   is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC   for treating or inhibiting development of cancer, e.g. lung cancer. This
CC   sequence represents a polynucleotide associated with the compositions and
CC   methods for the therapy and diagnosis of lung cancer
XX
SQ   Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.44e-58        Length:          822
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       7               Gaps:            0

US-09-684-215B-18 (1-128) x ACA03284 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 10
ABK27798
ID   ABK27798 standard; cDNA; 894 BP.
XX
AC   ABK27798;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human colon cancer expressed sequence tag, Seq ID no 236.
XX
KW   Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW   expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   WO200196390-A2.
XX
PD   20-DEC-2001.
XX
PF   08-JUN-2001; 2001WO-US018577.
XX
PR   09-JUN-2000; 2000US-0210821P.
PR   18-DEC-2000; 2000US-0256571P.
PR   10-MAY-2001; 2001US-0290240P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Jiang Y,  Hepler WT,  Clapper JD,  Wang A,  Secrist H;
XX
DR   WPI; 2002-139708/18.
DR   P-PSDB; AAU81532.
XX
PT   Novel isolated polynucleotide encoding a polypeptide comprising a portion
PT   of colon tumor protein, useful for detection, diagnosis and therapy of
PT   human colon cancer.
XX
PS   Claim 1; Page 215-216; 220pp; English.
XX
CC   The invention relates to an isolated polynucleotide (I) encoding a
CC   polypeptide (II) comprising at least a portion of a colon tumour protein.
CC   (I), (II) and antibody (III) to (II) are useful for determining the
CC   presence of a cancer in a patient. (I), (II) or antigen presenting cells
CC   expressing (I) is useful for stimulating and/or expanding T cells
CC   specific for a tumour protein, by contacting T cells with (I), (II) or
CC   antigen-presenting cells that express (I), under conditions and for a
CC   time sufficient to permit the stimulation and/or expansion of T cells.
CC   (I), (II), or antigen presenting cells that express (II) are useful for
CC   treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC   isolated from a patient with (I), (II) or antigen presenting cells that
CC   express (II), such that T cells proliferate, and administering to the
CC   patient an effective amount of the proliferated T cells, thus inhibiting
CC   the development of a cancer in the patient. (I) or (II) is useful in
CC   vaccines and pharmaceutical compositions for prevention and treatment of
CC   colon malignancies and for the diagnosis and monitoring of such cancers.
CC   (I), (II) or (III) is useful for detection, diagnosis and/or therapy of
CC   human colon cancer. (I) is useful as a probe or primer for nucleic acid
CC   hybridisation, and in the design and preparation of ribozyme molecules
CC   for inhibiting expression of (II) in tumour cells. ABK27564-ABK27807
CC   represent novel human colon cancer coding sequences and primers of the
CC   invention
XX
SQ   Sequence 894 BP; 207 A; 269 C; 237 G; 181 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.8e-58         Length:          894
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215B-18 (1-128) x ABK27798 (1-894)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 11
ABL49257
ID ABL49257 standard; cDNA; 900 BP.
XX
AC ABL49257;
XX
DT 01-MAY-2002 (first entry)
XX
DE Ra12/C-terminal portion of L763P fusion polynucleotide SEQ ID NO:353.
XX
KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW immune response; ss.
XX
OS Homo sapiens.
XX
PN WO200200174-A2.
XX
PD 03-JAN-2002.
XX
PF 28-JUN-2001; 2001WO-US021065.
XX
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI Mcneill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
DR WPI; 2002-090513/12.
DR P-PSDB; ABB75000.
XX
PT Polynucleotides encoding lung tumor polypeptides, useful for treating
PT lung cancer or stimulating an immune response.
XX
PS Claim 1; Page 333-334; 374pp; English.
XX
CC The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC proteins, T cell populations, or antigen presenting cells that express
CC the lung tumour proteins are useful for treating lung cancer or
CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC ABB75070 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 900 BP; 219 A; 239 C; 246 G; 196 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.83e-58 Length: 900
Score: 653.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-684-215B-18 (1-128) x ABL49257 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 12
ABQ92443
ID ABQ92443 standard; DNA; 900 BP.
XX
AC ABQ92443;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human lung cancer associated DNA sequence SEQ ID NO:353.
XX
KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200247534-A2.
XX
PD 20-JUN-2002.
XX

PF 30-NOV-2001; 2001WO-US047576.
XX
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR WPI; 2002-583465/62.
DR P-PSDB; ABP61920.
XX
PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX
PS Claim 8; Page 340-341; 381pp; English.
XX
CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridises to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotide to a
CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC vaccines. (I) is useful as a marker to indicate the presence or absence
CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC ABP61992 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 900 BP; 219 A; 239 C; 246 G; 196 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.83e-58	Length:	900
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-18 (1-128) x ABQ92443 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381
```

```
Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 13
ADA28443
ID ADA28443 standard; cDNA; 900 BP.
XX
AC ADA28443;
XX
DT 20-NOV-2003 (first entry)
XX
DE Lung tumour protein L763P C-terminal/Ra12 fusion DNA.
XX
KW cancer; lung cancer; gene therapy; vaccine; human;
KW lung squamous cell carcinoma; gene; ss; ra12.
XX
OS Homo sapiens.
XX
PN US2003064947-A1.
XX
PD 03-APR-2003.
XX
PF 30-NOV-2001; 2001US-00007700.
XX
PR 18-MAR-1998; 98US-00040802.
PR 27-JUL-1998; 98US-00123912.
PR 22-DEC-1998; 98US-00221107.
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00466396.
PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PR 04-APR-2000; 2000US-00542615.
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR WPI; 2003-540798/51.
DR P-PSDB; ADA28444.
XX
PT New isolated polynucleotides and polypeptides useful for diagnosing,
PT preventing and/or treating cancer, particularly lung cancer.
XX
PS Claim 8; Page 253; 296pp; English.
XX
CC The invention describes isolated polynucleotides and polypeptides useful
CC for diagnosing, preventing and/or treating cancer, particularly lung
CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC specification; complements of the nucleotide sequences cited above; at
CC least 10 contiguous resides of the nucleotide sequences cited above; a
CC sequence that hybridise to any of the nucleotide sequences under highly
CC stringent conditions; a sequence that is at least 75 or 90% identical to
CC the above nucleotide sequences; or degenerate variants of the above
CC nucleotide sequences. The composition and methods are useful in
CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC in gene therapy and in vaccines. This sequence encodes a fusion protein
CC of human lung tumour protein L673P C-terminal and Ra12.
XX
SQ Sequence 900 BP; 219 A; 239 C; 246 G; 196 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:              3.83e-58        Length:         900
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     8               Gaps:           0

US-09-684-215B-18 (1-128) x ADA28443 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 14
AAH93905
ID   AAH93905 standard; cDNA; 915 BP.
XX
AC   AAH93905;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Ra12-P775P-ORF3 construct cDNA sequence.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US001574.
XX
PR   14-JAN-2000; 2000US-00483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for diagnosing,
PT   monitoring and treating prostate cancer in a patient and for use in
PT   vaccines.
XX
PS   Claim 8; Page 497-498; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated T cells
CC   prepared using (I) or (II) are used treat cancer in a patient. (I) and
CC   the antibodies are also used in the detection of cancer in a patient. The
CC   cancer that is diagnosed or treated is particularly prostate cancer. (I)
CC   and (II) can be used in vaccines. The antibodies or (I) can be used for
CC   monitoring the progression of cancer in a patient. (I) and (II) can also
CC   be used to improve diagnostic and therapeutic methods for prostate
CC   cancer. They can indicate the level of metastasis as well as the prostate
CC   volume. AAH93357 to AAH93944 and AAM01115 to AAM01318 represent
CC   polynucleotide and amino acid sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.91e-58        Length:         915
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x AAH93905 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 15
AAS64141
ID   AAS64141 standard; cDNA; 915 BP.
XX
AC   AAS64141;
XX
DT   29-JAN-2002  (first entry)
XX
DE   Human /M. tuberculosis Ra12 fusion protein RA12-P775P-ORF3 cDNA.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
OS   Microbacterium; tuberculosis.
OS   Synthetic.
```

```
OS   Chimeric.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US009919.
XX
PR   27-MAR-2000; 2000US-00536857.
PR   09-MAY-2000; 2000US-00568100.
PR   12-MAY-2000; 2000US-00570737.
PR   13-JUN-2000; 2000US-00593793.
PR   27-JUN-2000; 2000US-00605783.
PR   09-AUG-2000; 2000US-00636215.
PR   29-AUG-2000; 2000US-00651236.
PR   06-SEP-2000; 2000US-00657279.
PR   02-OCT-2000; 2000US-00679426.
PR   10-OCT-2000; 2000US-00685166.
PR   09-NOV-2000; 2000US-00709729.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI   Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI   Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR   WPI; 2001-639232/73.
DR   P-PSDB; AAU69902.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer.
XX
PS   Example 17; Page 536-537; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific polynucleotides,
CC   polypeptides, fusion proteins of the polypeptides, antibodies raised
CC   against the polypeptides (or antigenic epitopes derived from them) and
CC   antigen-presenting cells expressing the polypeptides. The antibodies are
CC   useful for detecting the presence of cancer, especially prostate cancer.
CC   The polypeptides, polynucleotides and the antigen-presenting cells are
CC   useful for stimulating and/or expanding T cells specific for a tumour
CC   protein, and for inhibiting the development of cancer especially prostate
CC   cancer. Compositions comprising the polynucleotide and/or polypeptide are
CC   useful for stimulating an immune response, and for treating cancer. The
CC   oligonucleotide is useful for detecting cancer. The present sequence is a
CC   prostate specific polynucleotide of the invention
XX
SQ   Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                3.91e-58        Length:          915
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-18 (1-128) x AAS64141 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 16
ACA59949
ID   ACA59949 standard; cDNA; 915 BP.
XX
AC   ACA59949;
XX
DT   10-JUN-2003  (first entry)
XX
DE   Prostate cancer therapy associated cDNA #650.
XX
KW   Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;
KW   immunogen; cancer; prostate specific antigen; PSA;
KW   prostatic acid phosphatase; PAP; prostate specific membrane antigen;
KW   PSMA; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002192763-A1.
XX
PD   19-DEC-2002.
XX
PF   29-JUN-2001; 2001US-00895793.
XX
PR   04-OCT-1999;   99US-0157455P.
PR   04-OCT-2000; 2000US-00679272.
PR   28-MAR-2001; 2001US-00822827.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
PA   (HURA/) HURAL J.
PA   (MCNE/) MCNEILL P D.
PA   (HOUG/) HOUGHTON R L.
PA   (DBAS/) Y DE BASSOLS C V.
PA   (FOYT/) FOY T M.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI   Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI   Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;
PI   Mcneill PD, Houghton RL, Y De BassolsCV, Foy TM;
XX
DR   WPI; 2001-245062/25.
XX
PT   Prostate specific protein and its encoding polynucleotide, useful for the
PT   treatment and diagnosis of prostate cancer.
```

```
XX
PS   Example 17; SEQ ID NO 834; 85pp; English.
XX
CC   The invention describes a fusion protein comprising at least one amino
CC   acid sequence of immunogenic portions of any of the 3 sequences not
CC   defined in the specification, or sequences having at least 70 or 90 %
CC   sequence identity to any one of the 35 sequences defined in the USPTO web
CC   site, which is encoded by any of the 4 nucleotide sequences not defined
CC   in the specification. The fusion protein, composition and methods are
CC   useful for diagnosing, preventing and/or treating cancer, particularly
CC   prostate cancer. The proteins are useful as markers to indicate the
CC   presence or absence of cancer. This sequence represents a prostate cancer
CC   therapy associated cDNA. Note: The sequence data for this patent did not
CC   form part of the printed specification, but was obtained in electronic
CC   format directly from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=US20020192763
XX
SQ   Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.91e-58        Length:         915
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     5               Gaps:           0

US-09-684-215B-18 (1-128) x ACA59949 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 17
ABL95512
ID   ABL95512 standard; cDNA; 915 BP.
XX
AC   ABL95512;
XX
DT   29-AUG-2003  (revised)
DT   19-JUL-2002  (first entry)
XX
DE   Ra12- P775P-ORF3 construct cDNA sequence SEQ ID NO 834.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
XX
OS   Mycobacterium tuberculosis.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-00759143.
XX
PR   25-FEB-1997;   97US-00806099.
PR   01-AUG-1997;   97US-00904804.
PR   10-FEB-1998;   98US-00020956.
PR   25-FEB-1998;   98US-00030607.
PR   14-JUL-1998;   98US-00115453.
PR   23-SEP-1998;   98US-00159812.
PR   15-JAN-1999;   99US-00232149.
PR   09-APR-1999;   99US-00288946.
PR   13-JUL-1999;   99US-00352616.
PR   12-NOV-1999;   99US-00439313.
PR   18-NOV-1999;   99US-00443686.
PR   14-JAN-2000; 2000US-00483672.
PR   27-MAR-2000; 2000US-00536857.
PR   09-MAY-2000; 2000US-00568100.
PR   12-MAY-2000; 2000US-00570737.
PR   13-JUN-2000; 2000US-00593793.
PR   27-JUN-2000; 2000US-00605783.
PR   10-AUG-2000; 2000US-00636215.
PR   29-AUG-2000; 2000US-00651236.
PR   06-SEP-2000; 2000US-00657279.
PR   02-OCT-2000; 2000US-00679426.
PR   10-OCT-2000; 2000US-00685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer.
XX
PS   Example 17; SEQ ID NO 834; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention. (Updated on 29-AUG-2003 to standardise OS
CC   field)
XX
SQ   Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.91e-58        Length:         915
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
```

```
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     6              Gaps:           0

US-09-684-215B-18 (1-128) x ABL95512 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 18
ACC95676
ID   ACC95676 standard; cDNA; 915 BP.
XX
AC   ACC95676;
XX
DT   28-AUG-2003  (first entry)
XX
DE   Prostate tumour specific cDNA sequence SEQ ID 834.
XX
KW   Cytostatic; gene therapy; prostate-specific protein; PSP; human;
KW   immune response; prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200289747-A2.
XX
PD   14-NOV-2002.
XX
PF   09-MAY-2002; 2002WO-US014753.
XX
PR   09-MAY-2001; 2001US-00852911.
PR   29-JUN-2001; 2001US-00895814.
PR   10-DEC-2001; 2001US-00012896.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Henderson RA;
PI   Kalos MD,  Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS;
PI   Carter D,  Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Hural J;
PI   Mcneill PD,  Houghton RL,  Vinals Y De BassolsC,  Foy TM,  Watanabe Y;
PI   Deng T;
XX
DR   WPI; 2003-167130/16.
XX
PT   New prostate-specific proteins and genes, useful in gene therapy,
PT   particularly for stimulating an immune response in a patient, or treating
PT   prostate cancer in a patient, as well as for diagnosing prostate cancer
PT   in a patient.
XX
PS   Example 17; Page 605-606; 691pp; English.
XX
CC   The present invention relates to novel prostate-specific proteins (PSP)
CC   and their coding sequences. The PSPs and their coding sequences are
CC   useful for stimulating an immune response in a patient, or for treating
CC   prostate cancer in a patient and for determining, detecting or diagnosing
CC   the presence of a cancer in a patient. The present sequence was used to
CC   illustrate the invention
XX
SQ   Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.91e-58       Length:         915
Score:                  653.00         Matches:        128
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     7              Gaps:           0

US-09-684-215B-18 (1-128) x ACC95676 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 19
ADB14284
ID   ADB14284 standard; cDNA; 915 BP.
XX
AC   ADB14284;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human prostate specific protein P775P-ORF3-RA12 construct cDNA.
XX
KW   Human; ss; prostate specific cDNA; cytostatic; immunostimulant;
KW   gene therapy; cell therapy; vaccine; T-cell epitope;
KW   class I major histocompatibility complex allele; MHC; prostate cancer;
KW   tumour; antigen presenting cell; gene; fusion protein.
XX
OS   Chimeric.
OS   Homo sapiens.
OS   Mycobacterium tuberculosis.
XX
PN   US2003185830-A1.
XX
PD   02-OCT-2003.
XX
```

PF 12-NOV-2002; 2002US-00294025.
XX
PR 25-FEB-1997; 97US-00806099.
PR 01-AUG-1997; 97US-00904804.
PR 09-FEB-1998; 98US-00020956.
PR 25-FEB-1998; 98US-00030607.
PR 14-JUL-1998; 98US-00115453.
PR 23-SEP-1998; 98US-00159812.
PR 15-JAN-1999; 99US-00232149.
PR 09-APR-1999; 99US-00288946.
PR 13-JUL-1999; 99US-00352616.
PR 12-NOV-1999; 99US-00439313.
PR 18-NOV-1999; 99US-00443686.
PR 14-JAN-2000; 2000US-00483672.
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
PR 12-JAN-2001; 2001US-00759143.
PR 09-FEB-2001; 2001US-00780669.
PR 09-MAY-2001; 2001US-00852911.
PR 29-JUN-2001; 2001US-00895814.
PR 10-DEC-2001; 2001US-00012896.
PR 09-MAY-2002; 2002US-00144678.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Stolk JA, Kalos MD;
XX
DR WPI; 2003-756193/71.
DR P-PSDB; ADB14285.
XX
PT New isolated polypeptide for use in a vaccine for stimulating an immune
PT response, or for treating or diagnosis cancer, preferably prostate
PT cancer.
XX
PS Example 17; Page; 101pp; English.
XX
CC The invention relates to an isolated polypeptide comprising no more than
CC 11-542 amino acids of ADB13563 comprising a sequence ADB14487. The
CC peptides comprise a fragment ADB13563 of that contain naturally processed
CC T-cell epitopes for 3 class I major histocompatibility complex (MHC)
CC alleles. ADB13563 is a polypeptide encoded by a human prostate specific
CC cDNA, one of 648 disclosed as new. Also included are nucleic acids
CC encoding the proteins and peptides, expression vectors, a host cell
CC transformed with the vector, an isolated antibody (or antigen binding
CC fragment) that specifically binds to the protein or peptide, detecting
CC the presence of a cancer in a patient (comprising contacting a patient
CC sample with a binding agent that binds to the peptides or a polypeptide
CC appearing as ADB13558, detecting the amount of polypeptide that binds to
CC the agent and comparing the amount of polypeptide to a predetermined cut-
CC off value to determine the presence of cancer), a fusion protein
CC comprising the peptides or proteins, stimulating or expanding T cells
CC specific for a tumour protein comprising contacting T cells with the
CC peptides or the isolated T cell population, treating prostate cancer in a
CC patient comprising administering a composition comprising the peptides,
CC nucleic acids, antibodies or compounds, determining the presence of a
CC cancer in a patient and treating prostate cancer in a patient comprising
CC incubating cluster of differentiation (CD4+) and/or CD8+ T cells isolated
CC from a patient with the peptides or antigen presenting cells that express
CC (the peptides so that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The peptides (or an oligonucleotide
CC that hybridises to nucleic acid encoding them), is used to detect the
CC presence of cancer in a patient. The peptides, nucleic acids encoding, or
CC antigen-presenting cells expressing the nucleic acid, are used to
CC stimulate or expand T cells specific for a tumour protein. The peptides,
CC nucleic acids, antibodies, fusion proteins, T cell populations or antigen
CC presenting cells are used to stimulate an immune response or treat
CC prostate cancer in a patient. The present sequence is a cDNA encoding a
CC fusion protein comprising a prostate specific protein. Note: Except where
CC otherwise indicated, the sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030185830.
XX
SQ Sequence 915 BP; 167 A; 280 C; 284 G; 184 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.91e-58	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-18 (1-128) x ADB14284 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIlePro Ile 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 20
ABK39768
ID ABK39768 standard; cDNA; 945 BP.
XX
AC ABK39768;
XX
DT 21-MAY-2002 (first entry)
XX
DE DNA encoding lung tumour protein P801P ORF4 and Ra12 fusion protein.
XX
KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW gene; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200204514-A2.
XX
PD 17-JAN-2002.
XX
PF 10-JUL-2001; 2001WO-US022058.
XX
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.

PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, Mcnabb A;
PI Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
XX
DR WPI; 2002-164634/21.
DR P-PSDB; AAU85586.
XX
PT Novel polynucleotide encoding a lung tumor polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumor protein.
XX
PS Example 8; SEQ ID NO 1861; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide useful
CC for stimulating and/or expanding T cells specific for a tumour protein
CC for determining the presence of a cancer in a patient. A composition
CC containing the polynucleotide and/or polypeptide is useful for treating a
CC lung cancer in a patient. The polypeptide is useful for removing tumour
CC cells from a biological sample. The polynucleotide is also useful as
CC probe or primer to detect the level of mRNA encoding a tumour protein.
CC This sequence encodes a lung tumour associated protein or protein
CC fragment, described in the method of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 945 BP; 150 A; 318 C; 360 G; 117 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.06e-58	Length:	945
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-18 (1-128) x ABK39768 (1-945)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 21
ACA12097
ID ACA12097 standard; cDNA; 945 BP.
XX
AC ACA12097;
XX
DT 06-JUN-2003 (first entry)
XX
DE Human lung cancer protein L801P ORF4/Ra12 fusion protein cDNA.
XX
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8; RA12; gene.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2002197669-A1.
XX
PD 26-DEC-2002.
XX
PF 03-MAY-2001; 2001US-00849626.
XX
PR 13-DEC-2000; 2000US-00736457.
XX
PA (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;
PI Clapper JD;
XX
DR WPI; 2003-352750/33.
DR P-PSDB; ABU69561.
XX
PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
PT compositions, e.g. vaccines, for treating lung cancer.
XX
PS Example 8; Page; 72pp; English.
XX
CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC 90%, identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC in a patient, a fusion protein comprising the polypeptide, an
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells specific for a tumour
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4^+ and/or CD8^+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
CC presenting cells that express the polynucleotide, such that T cells
CC proliferate, administering to the patient an effective amount of the
CC proliferated T cells, and thus inhibiting the development of a cancer in
CC the patient. The polynucleotide, protein ands cells are useful in a
CC composition for stimulating an immune response in a patient, and for
CC treating a cancer in a patient (particularly lung cancer). The
CC oligonucleotide is useful for determining the presence of a cancer in a

CC patient. The protein and oligonucleotides are useful in pharmaceutical
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC or primer for nucleic acid hybridisation, and in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and proteins in tumour cells. An amplified portion of the
CC polynucleotide is useful for isolating a full-length gene from a suitable
CC library. The present sequence encodes a fusion protein of human RA12 with
CC the protein product of a cDNA (full length, extended or partial) isolated
CC from a library derived from lung tumour/cancer cells. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocId=20020197669
XX
SQ Sequence 945 BP; 150 A; 318 C; 360 G; 117 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.06e-58	Length:	945
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-09-684-215B-18 (1-128) x ACA12097 (1-945)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 22
ACA03283
ID ACA03283 standard; DNA; 945 BP.
XX
AC ACA03283;
XX
DT 22-MAY-2003 (first entry)
XX
DE Lung cancer therapyand diagnosis associated DNA #4.
XX
KW Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ds.
XX
OS Homo sapiens.
XX
PN US2002172952-A1.
XX
PD 21-NOV-2002.
XX
PF 10-JUL-2001; 2001US-00902941.
XX
PR 30-JUN-1999; 99US-00346492.
PR 15-OCT-1999; 99US-00419356.
PR 17-DEC-1999; 99US-00466867.
PR 30-DEC-1999; 99US-00476300.
PR 06-MAR-2000; 2000US-00519642.
PR 22-MAR-2000; 2000US-00533077.
PR 10-APR-2000; 2000US-00546259.
PR 27-APR-2000; 2000US-00560406.
PR 05-JUN-2000; 2000US-00589184.
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX
PA (CORI-) CORIXA CORP.
XX
PI Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;
PI Durham M, Carter D, Fanger GR, Vedvick TS, Bangur CS, Mcnabb A;
XX
DR WPI; 2003-328427/31.
XX
PT New polynucleotide, useful for preparing a composition for treating or
PT inhibiting development of cancer, e.g. lung cancer.
XX
PS Example 8; SEQID NO 1861; 82pp; English.
XX
CC The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC for treating or inhibiting development of cancer, e.g. lung cancer. This
CC sequence represents a polynucleotide associated with the compositions and
CC methods for the therapy and diagnosis of lung cancer
XX
SQ Sequence 945 BP; 150 A; 318 C; 360 G; 117 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.06e-58	Length:	945
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-09-684-215B-18 (1-128) x ACA03283 (1-945)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381
```

```
Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 23
ABL49256
ID   ABL49256 standard; cDNA; 1012 BP.
XX
AC   ABL49256;
XX
DT   01-MAY-2002  (first entry)
XX
DE   Ra12/N-terminal portion of L763P fusion polynucleotide SEQ ID NO:351.
XX
KW   Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW   immune response; ss.
XX
OS   Homo sapiens.
XX
PN   WO200200174-A2.
XX
PD   03-JAN-2002.
XX
PF   28-JUN-2001; 2001WO-US021065.
XX
PR   28-JUN-2000; 2000US-00606421.
PR   02-AUG-2000; 2000US-00630940.
PR   21-AUG-2000; 2000US-00643597.
PR   15-SEP-2000; 2000US-00662786.
PR   09-OCT-2000; 2000US-00685696.
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Wang A,  Skeiky YAW,  Li SX,  Kalos MD,  Henderson RA;
PI   Mcneill PD,  Fanger N,  Retter MW,  Marnerakis M,  Fanger GR;
PI   Vedvick TS,  Carter D,  Watanabe Y,  Peckham DW;
XX
DR   WPI; 2002-090513/12.
DR   P-PSDB; ABB74999.
XX
PT   Polynucleotides encoding lung tumor polypeptides, useful for treating
PT   lung cancer or stimulating an immune response.
XX
PS   Claim 1; Page 332; 374pp; English.
XX
CC   The present invention describes human lung tumour proteins. Human lung
CC   tumour proteins and polynucleotides have cytostatic and immunostimulant
CC   activities, and can be used in vaccine production. Compositions
CC   comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC   proteins, T cell populations, or antigen presenting cells that express
CC   the lung tumour proteins are useful for treating lung cancer or
CC   stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC   ABB75070 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 1012 BP; 235 A; 289 C; 289 G; 199 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.41e-58        Length:          1012
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-18 (1-128) x ABL49256 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 186

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 246

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 306

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 366

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 426

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        427 ACATTGGCCGAGGGACCCCCGGCC 450

RESULT 24
ABQ92442
ID   ABQ92442 standard; DNA; 1012 BP.
XX
AC   ABQ92442;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human lung cancer associated DNA sequence SEQ ID NO:351.
XX
KW   Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200247534-A2.
XX
PD   20-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-US047576.
XX
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
PR   28-JUN-2001; 2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Wang A,  Skeiky YAW,  Li SX,  Kalos MD,  Henderson RA;
PI   Mcneill PD,  Fanger N,  Retter MW,  Durham M,  Fanger GR,  Vedvick TS;
PI   Carter D,  Watanabe Y,  Peckham DW,  Cai F,  Foy TM;
XX
DR   WPI; 2002-583465/62.
DR   P-PSDB; ABP61919.
XX
PT   Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT   the polynucleotides, useful in pharmaceutical compositions such as
PT   vaccines and as markers to indicate the presence of lung cancer.
XX
PS   Claim 8; Page 339-340; 381pp; English.
XX
CC   The present invention describes isolated human lung carcinoma
CC   polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC   activity, and can be used in gene therapy and in vaccines. Compositions
CC   comprising (I) or (II) can be used for stimulating an immune response in
CC   a patient and for treating lung cancer in a patient. Oligonucleotides of
CC   (I) can be used for detecting the presence of a cancer in a patient, by
CC   obtaining a biological sample from the patient, contacting the biological
CC   sample with the oligonucleotide, detecting in the sample, an amount of
CC   polynucleotide that hybridises to the oligonucleotide and comparing the
```

```
CC   amount of polynucleotide that hybridises to the oligonucleotide to a
CC   predetermined cut-off value, and determining the presence of a cancer in
CC   the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC   vaccines. (I) is useful as a marker to indicate the presence or absence
CC   of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC   ABP61992 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 1012 BP; 235 A; 289 C; 289 G; 199 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.41e-58        Length:         1012
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x ABQ92442 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 186

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 246

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 306

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 366

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 426

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        427 ACATTGGCCGAGGGACCCCCGGCC 450

RESULT 25
ADA28441
ID   ADA28441 standard; cDNA; 1012 BP.
XX
AC   ADA28441;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Lung tumour protein L763P N-terminal/Ra12 fusion DNA.
XX
KW   cancer; lung cancer; gene therapy; vaccine; human;
KW   lung squamous cell carcinoma; gene; ss; ra12.
XX
OS   Homo sapiens.
XX
PN   US2003064947-A1.
XX
PD   03-APR-2003.
XX
PF   30-NOV-2001; 2001US-00007700.
XX
PR   18-MAR-1998;   98US-00040802.
PR   27-JUL-1998;   98US-00123912.
PR   22-DEC-1998;   98US-00221107.
PR   02-APR-1999;   99US-00285479.
PR   17-DEC-1999;   99US-00466396.
PR   30-DEC-1999;   99US-00476496.
PR   10-JAN-2000; 2000US-00480884.
PR   22-FEB-2000; 2000US-00510376.
PR   04-APR-2000; 2000US-00542615.
PR   28-JUN-2000; 2000US-00606421.
PR   02-AUG-2000; 2000US-00630940.
PR   21-AUG-2000; 2000US-00643597.
PR   15-SEP-2000; 2000US-00662786.
PR   09-OCT-2000; 2000US-00685696.
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
PR   28-JUN-2001; 2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Wang A,  Skeiky YAW,  Li SX,  Kalos MD,  Henderson RA;
PI   Mcneill PD,  Fanger N,  Retter MW,  Durham M,  Fanger GR,  Vedvick TS;
PI   Carter D,  Watanabe Y,  Peckham DW,  Cai F,  Foy TM;
XX
DR   WPI; 2003-540798/51.
DR   P-PSDB; ADA28441.
XX
PT   New isolated polynucleotides and polypeptides useful for diagnosing,
PT   preventing and/or treating cancer, particularly lung cancer.
XX
PS   Claim 8; Page 251-252; 296pp; English.
XX
CC   The invention describes isolated polynucleotides and polypeptides useful
CC   for diagnosing, preventing and/or treating cancer, particularly lung
CC   cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC   defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC   specification; complements of the nucleotide sequences cited above; at
CC   least 10 contiguous residues of the nucleotide sequences cited above; a
CC   sequence that hybridise to any of the nucleotide sequences under highly
CC   stringent conditions; a sequence that is at least 75 or 90% identical to
CC   the above nucleotide sequences; or degenerate variants of the above
CC   nucleotide sequences. The composition and methods are useful in
CC   diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC   in gene therapy and in vaccines. This sequence encodes a fusion protein
CC   of human lung tumour protein L673P N-terminal and Ra12.
XX
SQ   Sequence 1012 BP; 235 A; 289 C; 289 G; 199 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.41e-58        Length:         1012
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     8               Gaps:           0

US-09-684-215B-18 (1-128) x ADA28441 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 186

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 246

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 306

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 366
```

```
Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 426

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||||
Db        427 ACATTGGCCGAGGGACCCCCGGCC 450

RESULT 26
ABK69714
ID   ABK69714 standard; cDNA; 1035 BP.
XX
AC   ABK69714;
XX
DT   15-JUL-2002  (first entry)
XX
DE   DNA encoding human Ra12/WT1-E.
XX
KW   Human; mouse; cytostatic; immunostimulant; WT1; cancer; immune response;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200228414-A1.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US031139.
XX
PR   06-OCT-2000; 2000US-00684361.
PR   09-OCT-2000; 2000US-00685830.
PR   15-FEB-2001; 2001US-00785019.
PR   24-AUG-2001; 2001US-00938864.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Mcneill PD,  Smithgall M,  Moulton G,  Vedvick TS;
PI   Sleath PR,  Mossman S,  Evans L,  Spies AG,  Boydston J;
XX
DR   WPI; 2002-352217/38.
DR   P-PSDB; ABG33391.
XX
PT   Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
PT   treating and diagnosing cancer in a patient.
XX
PS   Claim 1; Page 240; 260pp; English.
XX
CC   The invention relates to an isolated WT1 polynucleotide (I) and
CC   polypeptide encoded by (I). The WT1 polynucleotides and polypeptides are
CC   used for treating and detecting cancer in a patient, and for stimulating
CC   an immune response in patient. ABK69648-ABK69724 represent WT1 coding
CC   sequences and PCR primers of the invention
XX
SQ   Sequence 1035 BP; 217 A; 355 C; 291 G; 172 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                4.53e-58        Length:          1035
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-18 (1-128) x ABK69714 (1-1035)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 63

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         64 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 123

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 183

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 243

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 303

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 363

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||||
Db        364 ACATTGGCCGAGGGACCCCCGGCC 387

RESULT 27
ABZ33699
ID   ABZ33699 standard; cDNA; 1035 BP.
XX
AC   ABZ33699;
XX
DT   30-JAN-2003  (first entry)
XX
DE   Human colon tumour cDNA for Ra12-C884P-PCRX2 SEQ ID NO:1084.
XX
KW   Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;
KW   tumour; immune response; immunostimulant; cytostatic; vaccine; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200283070-A2.
XX
PD   24-OCT-2002.
XX
PF   09-APR-2002; 2002WO-US011475.
XX
PR   10-APR-2001; 2001US-00833263.
PR   03-AUG-2001; 2001US-00922217.
PR   19-DEC-2001; 2001US-00025380.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Lodes MJ,  Secrist H,  Benson DR,  Meagher MJ,  Stolk JA;
PI   Wang T,  Jiang Y,  Smith CL,  King GE,  Wang A,  Clapper JD,  Skeiky YAW;
PI   Fanger GR,  Vedvick TS,  Carter D;
XX
DR   WPI; 2003-067548/06.
DR   P-PSDB; ABP55373.
XX
PT   New polynucleotide, useful for the preparation of a composition for
PT   stimulating an immune response against, or treating, cancer.
XX
PS   Claim 1; Page 474-475; 537pp; English.
XX
CC   The present invention describes compounds (I) for the immunotherapy and
CC   diagnosis of colon cancer. Also described: (1) a method for detecting the
CC   presence of cancer in a patient; (2) a method for stimulating and/or
CC   expanding T cells specific for a tumour protein; (3) an isolated T cell
CC   population comprising T cells prepared by the method of (2); (4) a method
CC   for stimulating an immune response in a patient; (5) a method for
CC   treating cancer in a patient; and (6) a method for inhibiting the
CC   development of cancer in a patient. (I) have immunostimulant and
CC   cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725
CC   and ABP55343 to ABP55391 represent human colon cancer/tumour related
CC   sequences used in the exemplification of the present invention
XX
```

```
SQ   Sequence 1035 BP; 252 A; 247 C; 274 G; 262 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.53e-58        Length:         1035
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     7               Gaps:           0

US-09-684-215B-18 (1-128) x ABZ33699 (1-1035)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 28
ADB67590
ID   ADB67590 standard; cDNA; 1035 BP.
XX
AC   ADB67590;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human RA12/WT1 E fusion protein cDNA.
XX
KW   Human; Wilm's tumour antigen; WT1; cytostatic; antigen; T cell;
KW   antigen-presenting cell; CD4+; CD8+; immune response; cancer; vaccine;
KW   leukaemia; antigen; ss; gene.
XX
OS   Chimeric.
OS   Homo sapiens.
OS   Mycobacterium tuberculosis.
XX
PN   US2003072767-A1.
XX
PD   17-APR-2003.
XX
PF   24-AUG-2001; 2001US-00938864.
XX
PR   30-SEP-1998;   98US-00164223.
PR   25-MAR-1999;   99US-00276484.
PR   06-OCT-2000; 2000US-00684361.
PR   09-OCT-2000; 2000US-00685830.
PR   15-FEB-2001; 2001US-00785019.
XX
PA   (GAIG/) GAIGER A.
PA   (MCNE/) MCNEILL P D.
PA   (SMIT/) SMITHGALL M.
PA   (MOUL/) MOULTON G.
PA   (VEDV/) VEDVICK T S.
PA   (SLEA/) SLEATH P R.
PA   (MOSS/) MOSSMAN S.
PA   (EVAN/) EVANS L.
PA   (SPIE/) SPIES A G.
PA   (BOYD/) BOYDSTON J.
XX
PI   Gaiger A,  Mcneill PD,  Smithgall M,  Moulton G,  Vedvick TS;
PI   Sleath PR,  Mossman S,  Evans L,  Spies AG,  Boydston J;
XX
DR   WPI; 2003-605957/57.
DR   P-PSDB; ADB67593.
XX
PT   Novel Wilms tumor polynucleotides encoding WT1 polypeptides, useful for
PT   detecting the presence of cancer in a patient, and in pharmaceutical
PT   compositions, e.g. vaccines, for treating cancers e.g. leukemia.
XX
PS   Claim 1; Page 145; 197pp; English.
XX
CC   The invention relates to an isolated Wilm tumour (WT1) polynucleotide
CC   comprising a full length protein, truncated protein, mutated protein or
CC   fusion protein. Also included are the encoded WT1 proteins, expression
CC   vectors, host cells, antibodies, detecting the presence of a cancer in a
CC   patient by contacting a biological sample with a binding agent that binds
CC   to a WT1 protein, an oligonucleotide that hybridises to a WT1
CC   polynucleotide, stimulating and/or expanding T cells specific for a
CC   tumour protein by contacting T cells with the WT1 polynucleotide, the WT1
CC   protein or antigen-presenting cells that express the WT1 protein, a
CC   composition (C1) (comprising physiologically acceptable carriers and
CC   immunostimulants as first component, and a second component selected from
CC   the WT1 polynucleotide, the WT1 protein, Ab, or T cells and antigen
CC   presenting cells that express the WT1 protein), inhibiting the
CC   development of a cancer in a patient (by incubating CD4+ and/or CD8+ T
CC   cells isolated from a patient with the WT1 polynucleotide, the WT1
CC   protein or antigen presenting cells that express the WT1 protein, such
CC   that T cell proliferate, administering the proliferated T cells to the
CC   patient, and thus inhibiting the development of a cancer in the patient)
CC   and a composition (C2) (comprising a WT1 polypeptide resuspended in a
CC   buffer comprising at least 1-3 sugars selected from trehalose, maltose,
CC   sucrose, fructose and glucose, at a concentration of 7-13%, and
CC   optionally ethanolamine, cysteine and Polysorbate-80, or WT1 polypeptide
CC   and MPL-SE or Enhanzyn). Also disclosed as anew are polypeptides
CC   comprising a variant of an immunogenic portion of WT1 polypeptide. C1 is
CC   useful for stimulating immune response in a patient, and for treating
CC   cancer in a patient. The oligonucleotide is also useful for determining
CC   the presence of a cancer in a patient. The WT1 polynucleotide and the WT1
CC   protein are useful in pharmaceutical compositions, e.g. vaccines. the WT1
CC   protein is useful as marker to indicate the presence or absence of a
CC   cancer. C1 is useful for inhibiting the development of a malignant
CC   disease in a patient, for preventing and treating metastatic diseases
CC   e.g. leukaemia and cancer, and for removing tumour cells from a
CC   biological sample. Ab (binding agent for the WT1 protein) is useful for
CC   detecting the presence of cancer in a patient. The present sequence
CC   encodes a Human WT1 fusion protein.
XX
SQ   Sequence 1035 BP; 217 A; 355 C; 291 G; 172 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.53e-58        Length:         1035
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-18 (1-128) x ADB67590 (1-1035)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 63

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 123

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 183

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        184 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 243

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 303

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 363

Qy        121 ThrLeuAlaGluGlyProProAla 128
                ||||||||||||||||||||||||
Db        364 ACATTGGCCGAGGGACCCCCGGCC 387

RESULT 29
ABK39775
ID   ABK39775 standard; cDNA; 1155 BP.
XX
AC   ABK39775;
XX
DT   21-MAY-2002  (first entry)
XX
DE   DNA encoding Ral2-L985P fusion protein.
XX
KW   Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW   gene; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200204514-A2.
XX
PD   17-JAN-2002.
XX
PF   10-JUL-2001; 2001WO-US022058.
XX
PR   11-JUL-2000; 2000US-00614124.
PR   29-AUG-2000; 2000US-00651563.
PR   08-SEP-2000; 2000US-00658824.
PR   26-SEP-2000; 2000US-00671325.
PR   06-OCT-2000; 2000US-00677419.
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Watanabe Y,  Henderson RA,  Johnson JC,  Retter MW;
PI   Marnerakis M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
PI   Wang A,  Fanger N,  Switzer A,  Mcneill PD,  Clapper JD;
XX
DR   WPI; 2002-164634/21.
DR   P-PSDB; AAU85593.
XX
PT   Novel polynucleotide encoding a lung tumor polypeptide useful for
PT   stimulating and/or expanding T cells specific for a tumor protein.
XX
PS   Claim 1; SEQ ID NO 1875; 223pp; English.
XX
CC   The invention describes an isolated polynucleotide and polypeptide useful
CC   for stimulating and/or expanding T cells specific for a tumour protein
CC   for determining the presence of a cancer in a patient. A composition
CC   containing the polynucleotide and/or polypeptide is useful for treating a
CC   lung cancer in a patient. The polypeptide is useful for removing tumour
CC   cells from a biological sample. The polynucleotide is also useful as
CC   probe or primer to detect the level of mRNA encoding a tumour protein.
CC   This sequence encodes a lung tumour associated protein or protein
CC   fragment, described in the method of the invention. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1155 BP; 239 A; 346 C; 302 G; 268 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                5.17e-58        Length:          1155
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-18 (1-128) x ABK39775 (1-1155)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
                ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 30
ACA12104
ID   ACA12104 standard; cDNA; 1155 BP.
XX
AC   ACA12104;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Human Ral2-L985P fusion protein cDNA.
XX
KW   Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW   T cell expansion; CD4; CD8; RA12; gene.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US2002197669-A1.
XX
PD   26-DEC-2002.
XX
PF   03-MAY-2001; 2001US-00849626.
XX
PR   13-DEC-2000; 2000US-00736457.
XX
PA   (BANG/) BANGUR C S.
```

PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;
PI Clapper JD;
XX
DR WPI; 2003-352750/33.
DR P-PSDB; ABU69568.
XX
PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
PT compositions, e.g. vaccines, for treating lung cancer.
XX
PS Example 10; Page; 72pp; English.
XX
CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC 90%, identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC in a patient, a fusion protein comprising the polypeptide, an
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells specific for a tumour
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4^+ and/or CD8^+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
CC presenting cells that express the polynucleotide, such that T cells
CC proliferate, administering to the patient an effective amount of the
CC proliferated T cells, and thus inhibiting the development of a cancer in
CC the patient. The polynucleotide, protein ands cells are useful in a
CC composition for stimulating an immune response in a patient, and for
CC treating a cancer in a patient (particularly lung cancer). The
CC oligonucleotide is useful for determining the presence of a cancer in a
CC patient. The protein and oligonucleotides are useful in pharmaceutical
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC or primer for nucleic acid hybridisation, and in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and proteins in tumour cells. An amplified portion of the
CC polynucleotide is useful for isolating a full-length gene from a suitable
CC library. The present sequence encodes a fusion protein of human RA12 with
CC the protein product of a cDNA (full length, extended or partial) isolated
CC from a library derived from lung tumour/cancer cells. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocId=20020197669
XX
SQ Sequence 1155 BP; 239 A; 346 C; 302 G; 268 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.17e-58	Length:	1155
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-09-684-215B-18 (1-128) x ACA12104 (1-1155)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

Search completed: April 29, 2004, 23:03:11
Job time : 368.096 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:32:52 ; Search time 3670.05 Seconds
(without alignments)
1511.672 Million cell updates/sec

Title: US-09-684-215B-18
Perfect score: 653
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....CTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13185/app_query.fasta_1.1180
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_6034_@runat_29042004_061304_13185 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	653	100.0	675	6	AR261272	AR261272 Sequence
2	653	100.0	675	6	AR400535	AR400535 Sequence
3	653	100.0	675	6	AR405802	AR405802 Sequence
4	653	100.0	675	6	AX201049	AX201049 Sequence
5	653	100.0	675	6	AX267848	AX267848 Sequence
6	653	100.0	822	6	AR277645	AR277645 Sequence
7	653	100.0	822	6	AX369152	AX369152 Sequence
8	653	100.0	894	6	AX351489	AX351489 Sequence
9	653	100.0	900	6	AR220690	AR220690 Sequence
10	653	100.0	900	6	AX365960	AX365960 Sequence
11	653	100.0	915	6	AR261281	AR261281 Sequence
12	653	100.0	915	6	AR400544	AR400544 Sequence
13	653	100.0	915	6	AR405811	AR405811 Sequence
14	653	100.0	915	6	AX201061	AX201061 Sequence
15	653	100.0	915	6	AX267860	AX267860 Sequence
16	653	100.0	945	6	AR277644	AR277644 Sequence
17	653	100.0	945	6	AX369151	AX369151 Sequence
18	653	100.0	1012	6	AR220689	AR220689 Sequence
19	653	100.0	1012	6	AX365958	AX365958 Sequence
20	653	100.0	1155	6	AX369165	AX369165 Sequence
21	653	100.0	1203	6	AR400556	AR400556 Sequence
22	653	100.0	1203	6	AR405823	AR405823 Sequence
23	653	100.0	1203	6	AX201078	AX201078 Sequence
24	653	100.0	1203	6	AX267877	AX267877 Sequence
25	653	100.0	1464	6	AR229410	AR229410 Sequence
26	653	100.0	1464	6	AR322028	AR322028 Sequence
27	653	100.0	1464	6	AX156105	AX156105 Sequence
28	653	100.0	1464	6	AX361955	AX361955 Sequence
29	653	100.0	1557	6	AR229398	AR229398 Sequence
30	653	100.0	1557	6	AR322016	AR322016 Sequence
31	653	100.0	1557	6	AX156089	AX156089 Sequence
32	653	100.0	1557	6	AX361939	AX361939 Sequence
33	653	100.0	1578	6	BD264515	BD264515 Compounds
34	653	100.0	1578	6	AR229328	AR229328 Sequence
35	653	100.0	1578	6	AR321946	AR321946 Sequence
36	653	100.0	1578	6	AX155945	AX155945 Sequence
37	653	100.0	1578	6	AX361795	AX361795 Sequence
38	653	100.0	1590	6	AX316986	AX316986 Sequence
39	653	100.0	1752	6	AR229413	AR229413 Sequence
40	653	100.0	1752	6	AR322031	AR322031 Sequence
41	653	100.0	1752	6	AX156109	AX156109 Sequence
42	653	100.0	1752	6	AX361959	AX361959 Sequence
43	653	100.0	1758	6	AR229401	AR229401 Sequence
44	653	100.0	1758	6	AR322019	AR322019 Sequence
45	653	100.0	1758	6	AX156093	AX156093 Sequence

ALIGNMENTS

RESULT 1

```
AR261272
LOCUS       AR261272                675 bp    DNA     linear   PAT 29-JAN-2003
DEFINITION  Sequence 822 from patent US 6321716.
ACCESSION   AR261272
VERSION     AR261272.1  GI:28072035
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Mashiki,Z. and Harada,J.
  TITLE     Negative pressure control apparatus for engine mounted in vehicle
  JOURNAL   Patent: US 6321716-A 822 27-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.4e-44        Length:          675
Score:                    653.00         Matches:         128
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-18 (1-128) x AR261272 (1-675)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy          41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy          61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy          81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy         101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy         121 ThrLeuAlaGluGlyProProAla 128
               ||||||||||||||||||||||||
Db         382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 2
AR400535
LOCUS       AR400535                675 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 822 from patent US 6620922.
ACCESSION   AR400535
VERSION     AR400535.1  GI:40144000
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6620922-A 822 16-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.4e-44        Length:          675
Score:                    653.00         Matches:         128
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-18 (1-128) x AR400535 (1-675)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy          41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy          61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy          81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy         101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy         121 ThrLeuAlaGluGlyProProAla 128
               ||||||||||||||||||||||||
Db         382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 3
AR405802
LOCUS       AR405802                675 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 822 from patent US 6630305.
ACCESSION   AR405802
VERSION     AR405802.1  GI:40154639
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6630305-A 822 07-OCT-2003;
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.4e-44        Length:          675
Score:                    653.00         Matches:         128
```

```
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-18 (1-128) x AR405802 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 4
AX201049
LOCUS       AX201049                675 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 679 from Patent WO0151633.
ACCESSION   AX201049
VERSION     AX201049.1  GI:15390857
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 679 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.4e-44        Length:        675
Score:                  653.00         Matches:       128
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-18 (1-128) x AX201049 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 5
AX267848
LOCUS       AX267848                675 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 822 from Patent WO0173032.
ACCESSION   AX267848
VERSION     AX267848.1  GI:16516494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 822 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.4e-44        Length:        675
Score:                  653.00         Matches:       128
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-18 (1-128) x AX267848 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 6
AR277645
LOCUS       AR277645                822 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 1862 from patent US 6509448.
ACCESSION   AR277645
VERSION     AR277645.1  GI:29711294
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
            Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6509448-A 1862 21-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.95e-44       Length:         822
Score:                    653.00         Matches:        128
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
Query Match:              100.00%        Indels:         0
DB:                       6              Gaps:           0

US-09-684-215B-18 (1-128) x AR277645 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 7
AX369152
LOCUS       AX369152                822 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 1862 from Patent WO0204514.
ACCESSION   AX369152
VERSION     AX369152.1  GI:18857170
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
            Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
            Mcnabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0204514-A 1862 17-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                2.95e-44       Length:         822
Score:                    653.00         Matches:        128
Percent Similarity:       100.00%        Conservative:   0
Best Local Similarity:    100.00%        Mismatches:     0
Query Match:              100.00%        Indels:         0
DB:                       6              Gaps:           0

US-09-684-215B-18 (1-128) x AX369152 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 8
AX351489
LOCUS       AX351489                894 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 236 from Patent WO0196390.
```

```
ACCESSION   AX351489
VERSION     AX351489.1  GI:18616835
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196390-A 236 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                3.21e-44        Length:         894
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-18 (1-128) x AX351489 (1-894)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 9
AR220690
LOCUS       AR220690                900 bp    DNA     linear   PAT 26-SEP-2002
DEFINITION  Sequence 353 from patent US 6426072.
ACCESSION   AR220690
VERSION     AR220690.1  GI:23327471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 900)
  AUTHORS   Wang,T., Fan,L., Kalos,M.D., Bangur,C.S., Hosken,N.A., Fanger,G.R.,
            Li,S.X., Wang,A., Skeiky,Y.A.W., Henderson,R.A. and McNeill,P.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6426072-A 353 30-JUL-2002;
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                3.24e-44        Length:         900
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-18 (1-128) x AR220690 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 10
AX365960
LOCUS       AX365960                900 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 353 from Patent WO0200174.
ACCESSION   AX365960
VERSION     AX365960.1  GI:18697458
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Wang,A., Skeiky,Y.A., Li,S.X., Kalos,M.D., Henderson,R.A.,
            Mcneill,P.D., Fanger,N., Retter,M.W., Marnerakis,M., Fanger,G.R.,
            Vedvick,T.S., Carter,D., Watanabe,Y. and Peckham,D.W.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0200174-A 353 03-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                3.24e-44        Length:         900
```

```
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AX365960 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 11
AR261281
LOCUS       AR261281                915 bp    DNA     linear   PAT 29-JAN-2003
DEFINITION  Sequence 834 from patent US 6321716.
ACCESSION   AR261281
VERSION     AR261281.1  GI:28072044
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Mashiki,Z. and Harada,J.
  TITLE     Negative pressure control apparatus for engine mounted in vehicle
  JOURNAL   Patent: US 6321716-A 834 27-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              3.29e-44        Length:         915
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AR261281 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 12
AR400544
LOCUS       AR400544                915 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 834 from patent US 6620922.
ACCESSION   AR400544
VERSION     AR400544.1  GI:40144016
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6620922-A 834 16-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              3.29e-44        Length:         915
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AR400544 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

```
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 13
AR405811
LOCUS       AR405811                915 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 834 from patent US 6630305.
ACCESSION   AR405811
VERSION     AR405811.1  GI:40154648
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6630305-A 834 07-OCT-2003;
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                3.29e-44        Length:         915
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-18 (1-128) x AR405811 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 14
AX201061
LOCUS       AX201061                915 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 691 from Patent WO0151633.
ACCESSION   AX201061
VERSION     AX201061.1  GI:15390868
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 691 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                3.29e-44        Length:         915
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-18 (1-128) x AX201061 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 15
AX267860
LOCUS       AX267860                915 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 834 from Patent WO0173032.
ACCESSION   AX267860
VERSION     AX267860.1  GI:16516503
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 834 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                3.29e-44        Length:         915
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-18 (1-128) x AX267860 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 16
AR277644
LOCUS       AR277644                 945 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 1861 from patent US 6509448.
ACCESSION   AR277644
VERSION     AR277644.1  GI:29711293
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 945)
  AUTHORS   Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
            Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6509448-A 1861 21-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .945
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                3.41e-44        Length:         945
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-18 (1-128) x AR277644 (1-945)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 17
AX369151
LOCUS       AX369151                 945 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 1861 from Patent WO0204514.
ACCESSION   AX369151
VERSION     AX369151.1  GI:18857169
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
            Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
            Mcnabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0204514-A 1861 17-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .945
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                3.41e-44        Length:         945
Score:                    653.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
```

```
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-684-215B-18 (1-128) x AX369151 (1-945)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 18
AR220689
LOCUS       AR220689               1012 bp    DNA     linear   PAT 26-SEP-2002
DEFINITION  Sequence 351 from patent US 6426072.
ACCESSION   AR220689
VERSION     AR220689.1  GI:23327470
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1012)
  AUTHORS   Wang,T., Fan,L., Kalos,M.D., Bangur,C.S., Hosken,N.A., Fanger,G.R.,
            Li,S.X., Wang,A., Skeiky,Y.A.W., Henderson,R.A. and McNeill,P.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6426072-A 351 30-JUL-2002;
FEATURES             Location/Qualifiers
     source          1. .1012
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              3.66e-44      Length:        1012
Score:                  653.00        Matches:       128
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-684-215B-18 (1-128) x AR220689 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 186

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 246

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 306

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 366

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 426

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        427 ACATTGGCCGAGGGACCCCCGGCC 450

RESULT 19
AX365958
LOCUS       AX365958               1012 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 351 from Patent WO0200174.
ACCESSION   AX365958
VERSION     AX365958.1  GI:18697457
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Wang,A., Skeiky,Y.A., Li,S.X., Kalos,M.D., Henderson,R.A.,
            Mcneill,P.D., Fanger,N., Retter,M.W., Marnerakis,M., Fanger,G.R.,
            Vedvick,T.S., Carter,D., Watanabe,Y. and Peckham,D.W.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0200174-A 351 03-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1012
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              3.66e-44      Length:        1012
Score:                  653.00        Matches:       128
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-684-215B-18 (1-128) x AX365958 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 186

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 246

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 306

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 366

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 426

Qy        121 ThrLeuAlaGluGlyProProAla 128
                |||||||||||||||||||||||
Db        427 ACATTGGCCGAGGGACCCCCGGCC 450

RESULT 20
AX369165
LOCUS       AX369165                1155 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 1875 from Patent WO0204514.
ACCESSION   AX369165
VERSION     AX369165.1  GI:18857176
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
            Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
            Mcnabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0204514-A 1875 17-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1155
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.2e-44       Length:          1155
Score:                    653.00        Matches:         128
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-18 (1-128) x AX369165 (1-1155)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
                |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 21
AR400556
LOCUS       AR400556                1203 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 851 from patent US 6620922.
ACCESSION   AR400556
VERSION     AR400556.1  GI:40144037
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1203)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6620922-A 851 16-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .1203
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                4.38e-44      Length:          1203
Score:                    653.00        Matches:         128
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-18 (1-128) x AR400556 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
                |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 22
AR405823
LOCUS       AR405823                1203 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 851 from patent US 6630305.
ACCESSION   AR405823
VERSION     AR405823.1  GI:40154660
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
```

```
            Unclassified.
REFERENCE   1  (bases 1 to 1203)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6630305-A 851 07-OCT-2003;
FEATURES             Location/Qualifiers
     source          1. .1203
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                4.38e-44       Length:          1203
Score:                    653.00         Matches:         128
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-18 (1-128) x AR405823 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 23
AX201078
LOCUS       AX201078               1203 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 708 from Patent WO0151633.
ACCESSION   AX201078
VERSION     AX201078.1  GI:15390883
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 708 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1203
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.38e-44       Length:          1203
Score:                    653.00         Matches:         128
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-18 (1-128) x AX201078 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 24
AX267877
LOCUS       AX267877               1203 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 851 from Patent WO0173032.
ACCESSION   AX267877
VERSION     AX267877.1  GI:16516515
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 851 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1203
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.38e-44       Length:          1203
```

```
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AX267877 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 25
AR229410
LOCUS       AR229410                1464 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 348 from patent US 6448234.
ACCESSION   AR229410
VERSION     AR229410.1  GI:27268663
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1464)
  AUTHORS   Fling,S.P.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: US 6448234-A 348 10-SEP-2002;
FEATURES             Location/Qualifiers
     source          1. .1464
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              5.37e-44        Length:         1464
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AR229410 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 26
AR322028
LOCUS       AR322028                1464 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 348 from patent US 6565856.
ACCESSION   AR322028
VERSION     AR322028.1  GI:33707372
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1464)
  AUTHORS   Skeiky,Y.A.W. and Scholler,J.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: US 6565856-A 348 20-MAY-2003;
FEATURES             Location/Qualifiers
     source          1. .1464
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              5.37e-44        Length:         1464
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AR322028 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321
```

```
Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 27
AX156105
LOCUS       AX156105               1464 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 348 from Patent WO0140474.
ACCESSION   AX156105
VERSION     AX156105.1  GI:14537130
KEYWORDS    .
SOURCE      Chlamydia trachomatis
  ORGANISM  Chlamydia trachomatis
            Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1
  AUTHORS   Probst,P., Bhatia,A., Skeiky,Y.A., Fling,S.P. and Scholler,J.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: WO 0140474-A 348 07-JUN-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1464
                     /organism="Chlamydia trachomatis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:813"
ORIGIN

Alignment Scores:
Pred. No.:                5.37e-44        Length:          1464
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-18 (1-128) x AX156105 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 28
AX361955
LOCUS       AX361955               1464 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 348 from Patent WO0208267.
ACCESSION   AX361955
VERSION     AX361955.1  GI:18694417
KEYWORDS    .
SOURCE      Chlamydia trachomatis
  ORGANISM  Chlamydia trachomatis
            Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1
  AUTHORS   Fling,S.P., Skeiky,Y.A., Probst,P. and Bhatia,A.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: WO 0208267-A 348 31-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1464
                     /organism="Chlamydia trachomatis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:813"
ORIGIN

Alignment Scores:
Pred. No.:                5.37e-44        Length:          1464
Score:                    653.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-18 (1-128) x AX361955 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 29
AR229398
LOCUS       AR229398               1557 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 332 from patent US 6448234.
ACCESSION   AR229398
VERSION     AR229398.1  GI:27268651
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1557)
  AUTHORS   Fling,S.P.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: US 6448234-A 332 10-SEP-2002;
FEATURES             Location/Qualifiers
```

```
     source          1. .1557
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              5.72e-44        Length:         1557
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AR229398 (1-1557)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 30
AR322016
LOCUS       AR322016                1557 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 332 from patent US 6565856.
ACCESSION   AR322016
VERSION     AR322016.1  GI:33707360
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1557)
  AUTHORS   Skeiky,Y.A.W. and Scholler,J.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: US 6565856-A 332 20-MAY-2003;
FEATURES             Location/Qualifiers
     source          1. .1557
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              5.72e-44        Length:         1557
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-18 (1-128) x AR322016 (1-1557)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

Search completed: April 30, 2004, 02:24:05
Job time : 3674.05 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:30:12 ; Search time 364.096 Seconds
(without alignments)
1493.479 Million cell updates/sec

Title: US-09-684-215B-23
Perfect score: 651
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....SVTWQTKSGGTRTGNVTLAE 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13179/app_query.fasta_1.1180
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_819_@runat_29042004_061304_13179 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	651	100.0	396	4	AAL40769	Aal40769 Nucleotid
2	651	100.0	672	4	AAL40772	Aal40772 Nucleotid
3	651	100.0	702	2	AAZ20206	Aaz20206 Mycobacte
4	651	100.0	702	4	AAL40770	Aal40770 Nucleotid
5	651	100.0	702	6	ABK14140	Abk14140 DNA encod
6	651	100.0	1002	6	AAD47078	Aad47078 Mycobacte
7	651	100.0	1002	6	AAD47077	Aad47077 Mycobacte
8	651	100.0	1002	6	AAD28336	Aad28336 Mycobacte
9	651	100.0	1002	6	AAD28337	Aad28337 Mycobacte
10	651	100.0	1068	2	AAX34251	Aax34251 Mycobacte
11	651	100.0	1143	2	AAX34252	Aax34252 Mycobacte
12	651	100.0	1629	5	AAS03793	Aas03793 M. tuberc
13	651	100.0	1742	4	AAL40771	Aal40771 Nucleotid
14	651	100.0	1871	4	AAL40768	Aal40768 Nucleotid
15	651	100.0	2181	8	ADA26353	Ada26353 Mycobacte
16	651	100.0	2190	6	AAD47084	Aad47084 Mycobacte
17	651	100.0	2190	6	AAD28343	Aad28343 Mycobacte
18	651	100.0	2191	4	AAL40773	Aal40773 Nucleotid
19	651	100.0	2286	6	ABK14128	Abk14128 DNA encod
20	651	100.0	2287	2	AAZ20194	Aaz20194 Mycobacte
21	651	100.0	2287	6	AAD47083	Aad47083 Mycobacte
22	651	100.0	2287	6	AAD28342	Aad28342 Mycobacte
23	651	100.0	2451	8	ADA26360	Ada26360 Mycobacte
24	651	100.0	2487	8	ADA26359	Ada26359 Mycobacte
25	651	100.0	2637	8	ADA26358	Ada26358 Mycobacte
26	651	100.0	2808	6	AAD47110	Aad47110 Mycobacte
27	651	100.0	2808	8	ADA26357	Ada26357 Mycobacte
28	651	100.0	3030	8	ADA26355	Ada26355 Mycobacte
29	651	100.0	3060	8	ADA26363	Ada26363 M. bovis
30	651	100.0	3104	8	ADA26362	Ada26362 Mycobacte
31	651	100.0	3474	8	ADA26361	Ada26361 Mycobacte
32	651	100.0	110000	4	AAI99682_01	Continuation (2 of
33	651	100.0	110000	4	AAI99683_01	Continuation (2 of
34	646	99.2	447	2	AAT91403	Aat91403 Mycobacte
35	646	99.2	447	2	AAT91466	Aat91466 Mycobacte
36	646	99.2	447	2	AAV44342	Aav44342 Mycobacte
37	646	99.2	447	2	AAV64450	Aav64450 M. tuberc
38	646	99.2	447	2	AAZ19040	Aaz19040 M. tuberc
39	646	99.2	447	2	AAZ19252	Aaz19252 M. tuberc
40	646	99.2	447	5	AAS03780	Aas03780 M. tuberc
41	646	99.2	447	6	AAD47080	Aad47080 Mycobacte
42	646	99.2	447	6	AAD28339	Aad28339 Mycobacte
43	646	99.2	1872	2	AAT91414	Aat91414 Mycobacte
44	646	99.2	1872	2	AAT91477	Aat91477 Mycobacte
45	646	99.2	1872	2	AAV44355	Aav44355 Mycobacte

ALIGNMENTS

RESULT 1
AAL40769
ID AAL40769 standard; DNA; 396 BP.
XX
AC AAL40769;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12 protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1. .396
FT /*tag= a
FT /product= "Ra12 protein"
FT /note= "No start or stop codon"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX

PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22138.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Claim 1; Fig 2; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12 protein
XX
SQ Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.69e-55	Length:	396
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-23 (1-128) x AAL40769 (1-396)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 120

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 180

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 240

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 300

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 360

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        361 ACAGGGAACGTGACATTGGCCGAG 384
```

RESULT 2
AAL40772
ID AAL40772 standard; DNA; 672 BP.
XX
AC AAL40772;
XX
DT 06-AUG-2003 (revised)
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12-mammaglobin fusion protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Mammalia.
OS Unidentified.
OS Chimeric.
XX

FH	Key	Location/Qualifiers
FT	CDS	4. .666
FT		/*tag= a
FT		/product= "Ra12-mammaglobin fusion protein"

XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22141.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Disclosure; Fig 5; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12-mammaglobin fusion protein. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.79e-54	Length:	672
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-23 (1-128) x AAL40772 (1-672)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84
```

```
Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
               |||||||||||||||||||||||
Db         385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 3
AAZ20206
ID AAZ20206 standard; DNA; 702 BP.
XX
AC AAZ20206;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.
XX
KW Tuberculosis; antigen; fusion protein; Mtb24; Ra12; DPPD; diagnosis;
KW therapy; vaccine; immunogen; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9951748-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Alderson M, Campos-Neto A;
XX
DR WPI; 1999-601610/51.
DR P-PSDB; AAY32071.
XX
PT New fusion proteins useful for diagnosis, prevention and treatment of
PT tuberculosis.
XX
PS Example; Fig 13A-B; 83pp; English.
XX
CC This DNA sequence includes a coding region for a recombinant
CC Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),
CC termed Mtb24, composed of the antigens Ra12 and DPPD. The DNA is useful
CC for the recombinant production of the fusion protein. Coding sequences
CC for the antigens were modified by PCR in order to facilitate their fusion
CC and subsequent expression of the fusion protein, and then ligated. The
CC invention provides fusion proteins (see AAY32059-71) containing at least
CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC encoding them are useful as vaccines for preventing tuberculosis
CC (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC progression, and treatment of tuberculosis. They are more effective
CC immunogens than mixtures of the individual protein components
XX
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.88e-54	Length:	702
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-684-215B-23 (1-128) x AAZ20206 (1-702)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
               |||||||||||||||||||||||
Db         385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 4
AAL40770
ID AAL40770 standard; DNA; 702 BP.
XX
AC AAL40770;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12-DPPD fusion protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 4. .696
FT /*tag= a
FT /product= "Ra12-DPPD fusion protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.

```
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22139.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Example 1; Fig 3; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12-DPPD fusion protein
XX
SQ   Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.88e-54        Length:          702
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       4               Gaps:            0

US-09-684-215B-23 (1-128) x AAL40770 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 5
ABK14140
ID   ABK14140 standard; DNA; 702 BP.
XX
AC   ABK14140;
XX
DT   29-AUG-2003  (revised)
DT   08-MAY-2002  (first entry)
XX
DE   DNA encoding antigenic fusion protein Ra12-DPPD (Mtb24).
XX
KW   Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW   tuberculostatic; immunogen; vaccine; Ra12-DPPD; Mtb24.
XX
OS   Mycobacterium tuberculosis.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .702
FT                   /*tag=  a
FT                   /product= "Mtb24 #1"
FT                   /partial
FT                   /transl_except= (pos:694. .696, aa:Xaa)
FT                   /note= "No start or stop codon. Xaa= In frame stop codon"
FT   CDS             2. .702
FT                   /*tag=  b
FT                   /product= "Mtb24 #2"
FT                   /partial
FT                   /note= "No start or stop codon"
FT                   /transl_except= (pos:263. .265, aa:Xaa) /transl_except=
FT                   (pos:353. .355, aa:Xaa) /transl_except= (pos:395. .397,
FT                   aa:Xaa) /transl_except= (pos:470. .472, aa:Xaa)
FT                   /transl_except= (pos:701. .702, aa:Ser)
FT                   /note= "This codon has an apparent 1 nucleotide deletion
FT                   which alters the reading frame. Xaa= In frame stop codon"
FT   CDS             3. .701
FT                   /*tag=  c
FT                   /product= "Mtb24 #3"
FT                   /partial
FT                   /note= "No start or stop codon"
FT                   /transl_except= (pos:1. .2, aa:Pro) /transl_except=
FT                   (pos:39. .41, aa:Xaa) /transl_except= (pos:321. .323,
FT                   aa:Xaa) /transl_except= (pos:339. .341, aa:Xaa)
FT                   /transl_except= (pos:450. .452, aa:Xaa) /transl_except=
FT                   (pos:621. .623, aa:Xaa)
FT                   /note= "No start or stop codon. Xaa= In frame stop codon"
XX
PN   US2002009459-A1.
XX
PD   24-JAN-2002.
XX
PF   07-APR-1999;   99US-00287849.
XX
PR   13-MAR-1997;   97US-00818112.
PR   01-OCT-1997;   97US-00942578.
PR   18-FEB-1998;   98US-00025197.
PR   07-APR-1998;   98US-00056556.
PR   30-DEC-1998;   98US-00223040.
XX
PA   (REED/) REED S G.
PA   (SKEI/) SKEIKY Y A.
PA   (DILL/) DILLON D C.
PA   (ALDE/) ALDERSON M.
PA   (CAMP/) CAMPOS-NETO A.
XX
PI   Reed SG,  Skeiky YA,  Dillon DC,  Alderson M,  Campos-Neto A;
XX
DR   WPI; 2002-171134/22.
DR   P-PSDB; AAU74600, AAU76541, AAU76542.
XX
PT   New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT   diagnosing, treating or preventing M. tuberculosis infection,
PT   particularly as vaccine for treating or preventing tuberculosis.
XX
PS   Example; Fig 13; 62pp; English.
XX
CC   The invention relates to a purified polypeptide which induces an immune
```

CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC fusion protein of the invention. This polynucleotide encodes 3 different
CC proteins, each in a different reading frame. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.88e-54	Length:	702
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x ABK14140 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 6
AAD47078
ID AAD47078 standard; DNA; 1002 BP.
XX
AC AAD47078;
XX
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium tuberculosis Ra35FLMutSA mutant antigenic protein DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; Ra35FLMutSA; mutant;
KW gene; antigen; ds.
XX
OS Mycobacterium tuberculosis.
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	CDS	4. .996
FT		/*tag= a
FT		/product= "Ra35FLMutSA mutant antigenic protein"

XX
PN WO200272792-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US008223.
XX
PR 13-MAR-2001; 2001US-0275837P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Brannon M, Guderian J;
XX
DR WPI; 2002-759844/82.
DR P-PSDB; AAE29703.
XX
PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT tuberculosis.
XX
PS Disclosure; Page 80-81; 155pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is M.
CC tuberculosis Ra35FLMutSA mutant antigenic protein encoding DNA
XX
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.85e-54	Length:	1002
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AAD47078 (1-1002)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
```

```
                    ||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAG 981

RESULT 7
AAD47077
ID   AAD47077 standard; DNA; 1002 BP.
XX
AC   AAD47077;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Mycobacterium tuberculosis mature Ra35 antigen encoding DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; Ra35 antigen; gene;
KW   ds.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .996
FT                   /*tag=  a
FT                   /product= "Ra35 mature antigenic protein"
FT                   /transl_except= (pos:547. .549, aa:Asp)
FT                   /transl_except= (pos:550. .552, aa:Ser)
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29702.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Disclosure; Page 79; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is M.
CC   tuberculosis mature Ra35 (N-terminus of MTB32A; RA35FL) antigen encoding
CC   DNA
XX
SQ   Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.85e-54        Length:          1002
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-23 (1-128) x AAD47077 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAG 981

RESULT 8
AAD28336
ID   AAD28336 standard; cDNA; 1002 BP.
XX
AC   AAD28336;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium sp. MTB32A (Ra35FL) mature protein encoding cDNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; MTB32A; Ra32FL; ss.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .996
FT                   /*tag=  a
FT                   /product= "Ra35 mature protein"
FT                   /transl_except= (pos:547. .549, aa:Asp)
FT                   /transl_except= (pos:550. .552, aa:Ser)
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Alderson M;
XX
DR   WPI; 2002-147798/19.
DR   P-PSDB; AAE17566.
XX
PT   Composition comprising MTB39 antigen and MTB32A antigen from
PT   Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS   Disclosure; Page 95; 136pp; English.
XX
CC   The present invention relates to fusion proteins containing at least two
```

CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a cDNA encoding
CC Mycobacterium species MTB32A (Ra32FL) mature protein
XX
SQ Sequence 1002 BP; 169 A; 331 C; 342 G; 160 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.85e-54	Length:	1002
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AAD28336 (1-1002)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAG 981
```

RESULT 9
AAD28337
ID AAD28337 standard; cDNA; 1002 BP.
XX
AC AAD28337;
XX
DT 22-APR-2002 (first entry)
XX
DE Mycobacterium species Ra35FLMutSA mutant cDNA.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; Ra32FLMutSA; mutant; mutein; ss.
XX
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .996
FT /*tag= a
FT /product= "Ra35FLMutSA protein"
XX
PN WO200198460-A2.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-US019959.
XX
PR 20-JUN-2000; 2000US-00597796.
PR 01-FEB-2001; 2001US-0265737P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Reed S, Alderson M;
XX
DR WPI; 2002-147798/19.
DR P-PSDB; AAE17567.
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Claim 67; Page 96-97; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a cDNA encoding
CC Mycobacterium species Ra32FL mature protein mutant, Ra32FLMutSA
XX
SQ Sequence 1002 BP; 169 A; 331 C; 341 G; 161 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.85e-54	Length:	1002
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AAD28337 (1-1002)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAG 981

RESULT 10
AAX34251
ID   AAX34251 standard; DNA; 1068 BP.
XX
AC   AAX34251;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Mycobacterium species nucleic acid sequence 50D.
XX
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS   Mycobacterium sp.
XX
PN   WO9909186-A2.
XX
PD   25-FEB-1999.
XX
PF   14-AUG-1998;   98WO-FR001813.
XX
PR   14-AUG-1997;   97FR-00010404.
PR   11-SEP-1997;   97FR-00011325.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Gicquel B,  Portnoie D,  Lim E,  Pelicic V,  Guigueno A;
PI   Goguet De La Salmoniere Y;
XX
DR   WPI; 1999-181045/15.
DR   P-PSDB; AAY05000.
XX
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in infection-associated
PT   protein expression.
XX
PS   Claim 22; Fig 50D; 309pp; French.
XX
CC   Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC   proteins from various Mycobacterium species microorganisms. The
CC   nucleotide sequences can be used as primers and probes for methods for
CC   detecting and identifying mycobacteria, especially belonging to the M.
CC   tuberculosis complex. The encoded proteins can be used in vaccines for
CC   immunisation against a bacterial or viral infection
XX
SQ   Sequence 1068 BP; 169 A; 352 C; 375 G; 172 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.06e-54      Length:         1068
Score:                  651.00        Matches:        128
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     2             Gaps:           0

US-09-684-215B-23 (1-128) x AAX34251 (1-1068)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAG 1053

RESULT 11
AAX34252
ID   AAX34252 standard; DNA; 1143 BP.
XX
AC   AAX34252;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Mycobacterium species nucleic acid sequence 50F.
XX
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection; ss.
XX
OS   Mycobacterium sp.
XX
PN   WO9909186-A2.
XX
PD   25-FEB-1999.
XX
PF   14-AUG-1998;   98WO-FR001813.
XX
PR   14-AUG-1997;   97FR-00010404.
PR   11-SEP-1997;   97FR-00011325.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Gicquel B,  Portnoie D,  Lim E,  Pelicic V,  Guigueno A;
PI   Goguet De La Salmoniere Y;
XX
DR   WPI; 1999-181045/15.
DR   P-PSDB; AAY04830.
XX
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in infection-associated
PT   protein expression.
XX
PS   Claim 22; Fig 50F; 309pp; French.
XX
CC   Sequences AAX34001-X34252 represent nucleic acids encoding secreted
CC   proteins from various Mycobacterium species microorganisms. The
CC   nucleotide sequences can be used as primers and probes for methods for
CC   detecting and identifying mycobacteria, especially belonging to the M.
CC   tuberculosis complex. The encoded proteins can be used in vaccines for
```

```
CC   immunisation against a bacterial or viral infection
XX
SQ   Sequence 1143 BP; 189 A; 373 C; 395 G; 186 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.31e-54        Length:         1143
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-09-684-215B-23 (1-128) x AAX34252 (1-1143)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAG 1128

RESULT 12
AAS03793
ID   AAS03793 standard; DNA; 1629 BP.
XX
AC   AAS03793;
XX
DT   29-AUG-2001  (first entry)
XX
DE   M. tuberculosis DNA encoding antigen HTCC#1 fusion protein #4.
XX
KW   TbRA12-HTCC#1; antigen; vaccine; tuberculosis; AIDS;
KW   acquired immunodeficiency disease; His Tag; ds.
XX
OS   Mycobacterium tuberculosis.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1629
FT                   /*tag=  a
FT                   /product= "TbRA12-HTCC#1"
FT                   /transl_except= (pos:1621. .1623,aa:Xaa)
FT                   /note= "Xaa= In frame STOP codon"
FT                   /partial
FT                   /note= "No start or stop codon. Although the sequence
FT                   does contain an in frame stop codon, 2 further amino
FT                   acids are shown in Figure 8 as being encoded by the
FT                   present sequence, without a further stop codon"
FT   misc_feature    25. .426
FT                   /*tag=  b
FT                   /note= "Region derived from TbRA12"
FT   misc_feature    427. .444
FT                   /*tag=  c
FT                   /note= "Region derived from Thrombin"
FT   misc_feature    445. .1629
FT                   /*tag=  d
FT                   /note= "Region derived from HTCC#1"
XX
PN   WO200124820-A1.
XX
PD   12-APR-2001.
XX
PF   10-OCT-2000; 2000WO-US028095.
XX
PR   07-OCT-1999;   99US-0158338P.
PR   07-OCT-1999;   99US-0158425P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Houghton RL,  Mcneill PD,  Dillon DC,  Lodes ML;
XX
DR   WPI; 2001-290576/30.
XX
PT   Vaccinating against Mycobacteria infections in mammals using fusion
PT   proteins comprising combinations of heterologous antigens.
XX
PS   Example 2; Fig 8; 168pp; English.
XX
CC   The sequence encodes Mycobacterium tuberculosis fusion protein, TbRA12-
CC   HTCC#1 and includes a His tag at the N-terminus to aid purification.
CC   Compositions comprising at least 2 heterologous antigens, as a fusion
CC   protein, and vectors expressing the fusion proteins are used as vaccines
CC   to prophylactically immunise mammals (especially humans) against
CC   infection by Mycobacteria. The compositions contain at least 2
CC   heterologous antigens that increase the serological sensitivity of
CC   individuals infected with tuberculosis, a disease frequently affecting
CC   patients with acquired immunodeficiency disease, AIDS
XX
SQ   Sequence 1629 BP; 298 A; 500 C; 535 G; 296 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              5e-54           Length:         1629
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     5               Gaps:           0

US-09-684-215B-23 (1-128) x AAS03793 (1-1629)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCCGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
```

```
                    ||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 13
AAL40771
ID AAL40771 standard; DNA; 1742 BP.
XX
AC AAL40771;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12-WT1 fusion protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
FH Key   Location/Qualifiers
FT CDS   4. .1740
FT   /*tag= a
FT   /product= "Ra12-WT1 fusion protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22140.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Disclosure; Fig 4; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12-WT1 fusion protein
XX
SQ Sequence 1742 BP; 372 A; 569 C; 490 G; 311 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.4e-54	Length:	1742
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-23 (1-128) x AAL40771 (1-1742)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 14
AAL40768
ID AAL40768 standard; DNA; 1871 BP.
XX
AC AAL40768;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding MTB32A protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Mycobacterium tuberculosis.
XX
FH Key   Location/Qualifiers
FT CDS   89. .1156
FT   /*tag= a
FT   /product= "Bacillus mycobacterium MTB32A protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22137.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Disclosure; Fig 1; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a

CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Mycobacterium tuberculosis MTB32A protein
XX
SQ Sequence 1871 BP; 317 A; 617 C; 605 G; 332 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.87e-54	Length:	1871
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-23 (1-128) x AAL40768 (1-1871)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

RESULT 15
ADA26353
ID ADA26353 standard; DNA; 2181 BP.
XX
AC ADA26353;
XX
DT 20-NOV-2003 (first entry)
XX
DE Mycobacterium MTB32-MTB39F fusion protein encoding DNA.
XX
KW ds; gene; fusion polypeptide; MTB32A; MTB39; antigen; MTB32A; MTB39;
KW MTB85A; tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS Chimeric.
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 4. .2175
FT /*tag= a
FT /product= "MTB32-MTB39F fusion polypeptide"
XX
PN WO2003070187-A2.
XX
PD 28-AUG-2003.
XX
PF 18-FEB-2003; 2003WO-US004903.
XX
PR 15-FEB-2002; 2002US-0357351P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J, Reed S;
XX
DR WPI; 2003-697554/66.
DR P-PSDB; ADA26354.
XX
PT New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT and MTB39 antigens, with or without the MTB85A antigen, from a
PT Mycobacterium species, useful for diagnosing, preventing and/or treating
PT tuberculosis infection.
XX
PS Claim 1; Fig 1; 112pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a fusion
CC polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC A polypeptide of the invention has tuberculostatic activity. A
CC polynucelotide of the invention may have a use in gene therapy, and as a
CC vaccine. The methods and compositions of the present invention are useful
CC for diagnosing, preventing and/or treating tuberculosis infection. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 2181 BP; 341 A; 693 C; 792 G; 355 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.01e-54	Length:	2181
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-09-684-215B-23 (1-128) x ADA26353 (1-2181)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
```

```
Db        958 ACAGGGAACGTGACATTGGCCGAG 981

RESULT 16
AAD47084
ID   AAD47084 standard; DNA; 2190 BP.
XX
AC   AAD47084;
XX
DT   29-AUG-2003  (revised)
DT   27-JAN-2003  (first entry)
XX
DE   Mycobacterium sp. MTB72FMutSA fusion protein encoding DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen;
KW   Ra35MutSA; Ra12; MTB72MutSA; chimeric; gene; ds.
XX
OS   Mycobacterium sp.
OS   Mycobacterium tuberculosis.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2190
FT                   /*tag= a
FT                   /product= "MTB72FMutSA fusion protein"
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29709.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Disclosure; Page 92-93; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is a
CC   DNA encoding MTB72F fusion protein. This fusion protein comprises
CC   Ra35MutSA mutant protein and Ra12 protein from M. tuberculosis and TbH9
CC   protein from Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS
CC   field)
XX
SQ   Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                7.05e-54        Length:          2190
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-23 (1-128) x AAD47084 (1-2190)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAG 405

RESULT 17
AAD28343
ID   AAD28343 standard; DNA; 2190 BP.
XX
AC   AAD28343;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium species MTB72FMutSA fusion protein encoding DNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; MTB72FMutSA; Ra12-TbH9-Ra35MutSA;
KW   mutant; mutein; ds.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2190
FT                   /*tag= a
FT                   /product= "MTB72FMutSA fusion protein"
FT   misc_feature    22. .417
FT                   /*tag= b
FT                   /note= "Ra12 DNA fragment"
FT   misc_feature    424. .1596
FT                   /*tag= c
FT                   /note= "TbH9FL DNA fragment"
FT   misc_feature    1603. .2187
FT                   /*tag= d
FT                   /note= "Ra35 DNA fragment"
FT   mutation        replace(2128, T)
FT                   /*tag= e
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
```

XX
PI Skeiky Y, Reed S, Alderson M;
XX
DR WPI; 2002-147798/19.
DR P-PSDB; AAE17573.
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Claim 81; Page 108-109; 136pp; English.
XX
CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is a DNA encoding
CC Mycobacterium species MTB72FMutSA (Ra12-TbH9-Ra35MutSA) mutant fusion
CC protein
XX
SQ Sequence 2190 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.05e-54	Length:	2190
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x AAD28343 (1-2190)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAG 405
```

RESULT 18
AAL40773
ID AAL40773 standard; DNA; 2191 BP.
XX
AC AAL40773;
XX
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12-H9-32A fusion protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1. .2190
FT /*tag= a
FT /product= "Ra12-H9-32A fusion protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX
PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22142.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Disclosure; Fig 6; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12-H9-32A fusion protein
XX
SQ Sequence 2191 BP; 340 A; 697 C; 797 G; 356 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	7.05e-54	Length:	2191
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-23 (1-128) x AAL40773 (1-2191)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAG 405
```

RESULT 19
ABK14128
ID ABK14128 standard; DNA; 2286 BP.
XX
AC ABK14128;
XX
DT 29-AUG-2003 (revised)
DT 08-MAY-2002 (first entry)
XX
DE DNA encoding antigenic fusion protein Ra12-TbH9-Ra35 (Mtb32-Mtb39).
XX
KW Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW tuberculostatic; immunogen; vaccine; Mtb32-Mtb39; Ra12; TbH9; Ra35;
KW Ra12-TbH9-Ra35.
XX
OS Mycobacterium tuberculosis.
OS Chimeric.
XX

FH	Key	Location/Qualifiers
FT	CDS	42. .2231
FT		/*tag= a
FT		/product= "Mtb32-Mtb39"
FT		/transl_except= (pos:498. .506, aa:Asn-Ala)
FT		/transl_except= (pos:597. .605, aa:Ala-Gln)
FT		/transl_except= (pos:798. .802, aa:Ala)
FT		/note= "This codon has an apparent 2 nucleotide insertion
FT		which alters the reading frame"

XX
PN US2002009459-A1.
XX
PD 24-JAN-2002.
XX
PF 07-APR-1999; 99US-00287849.
XX
PR 13-MAR-1997; 97US-00818112.
PR 01-OCT-1997; 97US-00942578.
PR 18-FEB-1998; 98US-00025197.
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (REED/) REED S G.
PA (SKEI/) SKEIKY Y A.
PA (DILL/) DILLON D C.
PA (ALDE/) ALDERSON M.
PA (CAMP/) CAMPOS-NETO A.
XX
PI Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX
DR WPI; 2002-171134/22.
DR P-PSDB; AAU74588.
XX
PT New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT diagnosing, treating or preventing M. tuberculosis infection,
PT particularly as vaccine for treating or preventing tuberculosis.
XX
PS Example; Fig 1; 62pp; English.
XX
CC The invention relates to a purified polypeptide which induces an immune
CC response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC useful for diagnosing, treating or preventing M. tuberculosis infection,
CC particularly tuberculosis infection. In particular, the polypeptides are
CC useful as a vaccine formulation with an adjuvant to afford long-term
CC protection in animals against the development of tuberculosis. The
CC protein coding sequence may be used to encode a protein product for use
CC as an immunogen to induce and/or enhance an immune response to M.
CC tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC fusion protein of the invention. (Updated on 29-AUG-2003 to standardise
CC OS field)
XX
SQ Sequence 2286 BP; 373 A; 719 C; 807 G; 381 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.:	7.41e-54	Length:	2286
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-23 (1-128) x ABK14128 (1-2286)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446
```

RESULT 20
AAZ20194
ID AAZ20194 standard; DNA; 2287 BP.
XX
AC AAZ20194;
XX
DT 17-JAN-2000 (first entry)
XX
DE Mycobacterium tuberculosis antigen fusion protein Mtb32A DNA.

XX
KW Tuberculosis; antigen; fusion protein; Mtb32A; Ra12; TbH9; Ra35;
KW diagnosis; therapy; vaccine; immunogen; ss.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT CDS 42. .2231
FT /*tag= a
XX
PN WO9951748-A2.
XX
PD 14-OCT-1999.
XX
PF 07-APR-1999; 99WO-US007717.
XX
PR 07-APR-1998; 98US-00056556.
PR 30-DEC-1998; 98US-00223040.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Alderson M, Campos-Neto A;
XX
DR WPI; 1999-601610/51.
DR P-PSDB; AAY32059.
XX
PT New fusion proteins useful for diagnosis, prevention and treatment of
PT tuberculosis.
XX
PS Example; Fig 1A-B; 83pp; English.
XX
CC This DNA sequence includes a coding region for a recombinant
CC Mycobacterium tuberculosis tri-antigen fusion protein (see AAY32059),
CC termed Mtb32A, composed of the antigens Ra12, TbH9 and Ra35. The DNA is
CC useful for the recombinant production of the fusion protein. Coding
CC sequences for the antigens were modified by PCR in order to facilitate
CC their fusion and subsequent expression of the fusion protein. 3 Coding
CC sequences for Ra12, TbH9 and Ra25 were ligated to encode Mtb32A. The
CC invention provides fusion proteins (see AAY32059-71) containing at least
CC 2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC encoding them are useful as vaccines for preventing tuberculosis
CC (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC progression, and treatment of tuberculosis. They are more effective
CC immunogens than mixtures of the individual protein components
XX
SQ Sequence 2287 BP; 374 A; 723 C; 805 G; 382 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	7.41e-54	Length:	2287
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-684-215B-23 (1-128) x AAZ20194 (1-2287)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446
```

RESULT 21
AAD47083
ID AAD47083 standard; DNA; 2287 BP.
XX
AC AAD47083;
XX
DT 29-AUG-2003 (revised)
DT 27-JAN-2003 (first entry)
XX
DE Mycobacterium sp. MTB72F fusion protein encoding DNA.
XX
KW Vaccine; immunity; diagnostic agent; gene therapy; TbH9; antigen; Ra35;
KW Ra12; MTB72F; chimeric; gene; ds.
XX
OS Mycobacterium sp.
OS Mycobacterium tuberculosis.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 42. .2231
FT /*tag= a
FT /product= "MTB72F fusion protein"
XX
PN WO200272792-A2.
XX
PD 19-SEP-2002.
XX
PF 13-MAR-2002; 2002WO-US008223.
XX
PR 13-MAR-2001; 2001US-0275837P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Brannon M, Guderian J;
XX
DR WPI; 2002-759844/82.
DR P-PSDB; AAE29708.
XX
PT New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT tuberculosis.
XX
PS Disclosure; Page 87-90; 155pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC polynucleotide sequence encoding an antigen or an antigenic fragment from
CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is a
CC DNA encoding MTB72F fusion protein. This fusion protein comprises Ra12
CC and Ra35 protein from Mycobacterium tuberculosis and TbH9 protein from
CC Mycobacterium sp. (Updated on 29-AUG-2003 to standardise OS field)

```
XX
SQ   Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              7.41e-54        Length:         2287
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-23 (1-128) x AAD47083 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 22
AAD28342
ID   AAD28342 standard; DNA; 2287 BP.
XX
AC   AAD28342;
XX
DT   22-APR-2002  (first entry)
XX
DE   Mycobacterium species MTB72F fusion protein encoding DNA.
XX
KW   Fusion protein; antigen; serological sensitivity; immune response;
KW   tuberculosis; infection; vaccine; MTB72F; Ra12-TbH9-Ra35; ds.
XX
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             42. .2231
FT                   /*tag=  a
FT                   /product= "MTB72F fusion protein"
FT   misc_feature    63. .458
FT                   /*tag=  b
FT                   /note= "Ra12 DNA fragment"
FT   misc_feature    465. .1637
FT                   /*tag=  c
FT                   /note= "TbH9FL DNA fragment"
FT   misc_feature    1644. .2228
FT                   /*tag=  d
FT                   /note= "Ra35 DNA fragment"
XX
PN   WO200198460-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019959.
XX
PR   20-JUN-2000; 2000US-00597796.
PR   01-FEB-2001; 2001US-0265737P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Alderson M;
XX
DR   WPI; 2002-147798/19.
DR   P-PSDB; AAE17572.
XX
PT   Composition comprising MTB39 antigen and MTB32A antigen from
PT   Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS   Claim 62; Page 103-106; 136pp; English.
XX
CC   The present invention relates to fusion proteins containing at least two
CC   Mycobacterium species antigens, nucleotides encoding them and
CC   compositions comprising such fusion proteins. The present invention
CC   particularly relates to nucleic acids encoding fusion proteins that
CC   include two or more individual M. tuberculosis antigens which increase
CC   the serological sensitivity of sera from individuals infected with
CC   tuberculosis and methods for their use in diagnosis, prevention and
CC   treatment of tuberculosis infection. Sequences of the invention are
CC   useful for eliciting an immune response in a mammal, e.g., human,
CC   immunised with BCG. They are useful in the diagnosis, treatment and
CC   prevention of Mycobacterium infection. The fusion proteins and the
CC   polynucleotides are useful as diagnostic tools in patients infected with
CC   Mycobacterium, in vitro and in vivo assays for detecting humoral
CC   antibodies or cell-mediated immunity against M. tuberculosis, for the
CC   diagnosis of an infection or monitoring of disease progression, as
CC   immunogens to generate or elicit a protective immune response in a
CC   patient and for raising anti-M. tuberculosis antibodies in a non-human
CC   animal. Sequences of the invention are also used as vaccines. MTB32A
CC   fusion proteins of the invention are useful as in vivo diagnostic agents
CC   for intradermal skin test. The present sequence is a DNA encoding
CC   Mycobacterium species MTB72F (Ra12-TbH9-Ra35) fusion protein
XX
SQ   Sequence 2287 BP; 372 A; 717 C; 814 G; 381 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.:              7.41e-54        Length:         2287
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-23 (1-128) x AAD28342 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
                ||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 23
ADA26360
ID   ADA26360 standard; DNA; 2451 BP.
XX
AC   ADA26360;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-DPV (fusion MTB81F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2445
FT                   /*tag=  a
FT                   /product= " MTB72F-DPV (fusion MTB81F) protein"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26367.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 8; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 2451 BP; 389 A; 785 C; 870 G; 407 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              8.03e-54        Length:         2451
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     8               Gaps:           0

US-09-684-215B-23 (1-128) x ADA26360 (1-2451)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 24
ADA26359
ID   ADA26359 standard; DNA; 2487 BP.
XX
AC   ADA26359;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-MTI (fusion MTB83F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2481
FT                   /*tag=  a
FT                   /product= "MTB72F-MTI (fusion MTB83F) protein"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26366.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 7; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
```

```
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 2487 BP; 394 A; 784 C; 896 G; 413 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.17e-54        Length:          2487
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       8               Gaps:            0

US-09-684-215B-23 (1-128) x ADA26359 (1-2487)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 25
ADA26358
ID   ADA26358 standard; DNA; 2637 BP.
XX
AC   ADA26358;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-Erd14 (fusion MTB89F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2631
FT                   /*tag=  a
FT                   /product= "MTB72F-Erd14 (fusion MTB89F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26365.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 6; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 2637 BP; 428 A; 840 C; 928 G; 441 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                8.74e-54        Length:          2637
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       8               Gaps:            0

US-09-684-215B-23 (1-128) x ADA26358 (1-2637)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 26
AAD47110
ID   AAD47110 standard; DNA; 2808 BP.
XX
AC   AAD47110;
XX
DT   29-AUG-2003  (revised)
DT   27-JAN-2003  (first entry)
XX
```

```
DE   Mycobacterium sp. MTB72F-Leishmania sp. MAPS (aka r95f)-fusion DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; MTB72F; MAPS;
KW   chimeric; gene; ds.
XX
OS   Mycobacterium sp.
OS   Leishmania sp.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2796
FT                   /*tag=  a
FT                   /product= "Mycobacterium sp. MTB72F-Leishmania sp. MAPS
FT                   (aka r95f) fusion protein"
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29731.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Example 6; Page 128-129; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
CC   polypeptide or its fragment. The Leishmania polynucleotide is selected
CC   from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC   are used in methods for eliciting immune response in mammals. They are
CC   useful as vaccines to elicit protective immunity against pathogenic
CC   microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC   polypeptides are used for enhancing the expression of polynucleotides, as
CC   in vivo diagnostic agents and for raising antibodies in a non-human
CC   animal. The invention is used in gene therapy. The present sequence is
CC   Mycobacterium sp. MTB72F-Leishmania sp. thiol specific antioxidant [TSA;
CC   MAPS (aka r95f)] fusion DNA. This sequence comprises Mycobacterium sp.
CC   MTB72F (a 72 kDa poly-protein fusion construct comprising Ra12-TbH9-Ra35)
CC   linked to the Leishmania TSA (MAPS). (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              9.4e-54       Length:          2808
Score:                  651.00        Matches:         128
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     6             Gaps:            0

US-09-684-215B-23 (1-128) x AAD47110 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 27
ADA26357
ID   ADA26357 standard; DNA; 2808 BP.
XX
AC   ADA26357;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-MAPS (fusion r95F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .2796
FT                   /*tag=  a
FT                   /product= "MTB72F-MAPS (fusion r95F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26364.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Disclosure; Fig 5; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
```

```
XX
SQ   Sequence 2808 BP; 466 A; 889 C; 982 G; 471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              9.4e-54         Length:         2808
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     8               Gaps:           0

US-09-684-215B-23 (1-128) x ADA26357 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 28
ADA26355
ID   ADA26355 standard; DNA; 3030 BP.
XX
AC   ADA26355;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB-102F fusion protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .3030
FT                   /*tag= a
FT                   /product= "MTB-120F fusion protein"
FT                   /note= "No stop codon given"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26356.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 5; Fig 3; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3030 BP; 498 A; 966 C; 1062 G; 504 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.03e-53        Length:         3030
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     8               Gaps:           0

US-09-684-215B-23 (1-128) x ADA26355 (1-3030)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 654

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 714

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 774

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 834

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 894

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 954

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        955 ACAGGGAACGTGACATTGGCCGAG 978

RESULT 29
ADA26363
ID   ADA26363 standard; DNA; 3060 BP.
XX
AC   ADA26363;
XX
DT   20-NOV-2003  (first entry)
XX
DE   M. bovis MTB72F and 85b complex antigen (fusion MTB103F) DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine; MTB72F; 85B.
```

```
XX
OS   Chimeric.
OS   Mycobacterium bovis.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .3054
FT                   /*tag= a
FT                   /product= "MTB72F and 85b complex (fusion MTB103F)"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26370.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 11; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3060 BP; 503 A; 977 C; 1070 G; 510 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.04e-53        Length:          3060
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       8               Gaps:            0

US-09-684-215B-23 (1-128) x ADA26363 (1-3060)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 30
ADA26362
ID   ADA26362 standard; DNA; 3104 BP.
XX
AC   ADA26362;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Mycobacterium MTB72F-hTCC#1 (fusion MTB102tm2F) protein encoding DNA.
XX
KW   ds; gene; fusion protein; MTB32A; MTB39; antigen; MTB32A; MTB39; MTB85A;
KW   tuberculosis; tuberculostatic; gene therapy; vaccine.
XX
OS   Chimeric.
OS   Mycobacterium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .3072
FT                   /*tag= a
FT                   /product= "MTB72F-hTCC#1 (fusion MTB102tm2F) protein"
XX
PN   WO2003070187-A2.
XX
PD   28-AUG-2003.
XX
PF   18-FEB-2003; 2003WO-US004903.
XX
PR   15-FEB-2002; 2002US-0357351P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J,  Reed S;
XX
DR   WPI; 2003-697554/66.
DR   P-PSDB; ADA26369.
XX
PT   New isolated nucleic acid encoding a fusion polypeptide with the MTB32A
PT   and MTB39 antigens, with or without the MTB85A antigen, from a
PT   Mycobacterium species, useful for diagnosing, preventing and/or treating
PT   tuberculosis infection.
XX
PS   Claim 84; Fig 10; 112pp; English.
XX
CC   The invention relates to a novel nucleic acid encoding a fusion
CC   polypeptide, comprising a MTB32A and MTB39 antigen, or MTB32A, MTB39 and
CC   MTB85A antigen from a Mycobacterium species of the tuberculosis complex.
CC   A polypeptide of the invention has tuberculostatic activity. A
CC   polynucelotide of the invention may have a use in gene therapy, and as a
CC   vaccine. The methods and compositions of the present invention are useful
CC   for diagnosing, preventing and/or treating tuberculosis infection. The
CC   present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 3104 BP; 510 A; 968 C; 1104 G; 522 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.06e-53        Length:          3104
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       8               Gaps:            0

US-09-684-215B-23 (1-128) x ADA26362 (1-3104)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

Search completed: April 29, 2004, 23:03:15
Job time : 368.096 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:43:07 ; Search time 469.129 Seconds
(without alignments)
1234.209 Million cell updates/sec

Title: US-09-684-215B-23
Perfect score: 651
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....SVTWQTKSGGTRTGNVTLAE 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13266/app_query.fasta_1.1180
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09684215_@CGN_1_1_370_@runat_29042004_061306_13266
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	651	100.0	702	9	US-09-287-849-27	Sequence 27, Appl
2	651	100.0	702	15	US-10-359-460-27	Sequence 27, Appl
3	651	100.0	1002	15	US-10-098-732A-3	Sequence 3, Appli
4	651	100.0	1002	15	US-10-098-732A-5	Sequence 5, Appli
5	651	100.0	1068	9	US-09-712-363-15	Sequence 15, Appl
6	651	100.0	2181	16	US-10-369-983-1	Sequence 1, Appli
7	651	100.0	2190	15	US-10-098-732A-17	Sequence 17, Appl
8	651	100.0	2287	9	US-09-287-849-1	Sequence 1, Appli
9	651	100.0	2287	15	US-10-359-460-1	Sequence 1, Appli
10	651	100.0	2287	15	US-10-098-732A-15	Sequence 15, Appl
11	651	100.0	2287	16	US-10-359-459-1	Sequence 1, Appli
12	651	100.0	2451	16	US-10-369-983-8	Sequence 8, Appli
13	651	100.0	2487	16	US-10-369-983-7	Sequence 7, Appli
14	651	100.0	2637	16	US-10-369-983-6	Sequence 6, Appli
15	651	100.0	2808	15	US-10-098-732A-64	Sequence 64, Appl
16	651	100.0	2808	16	US-10-369-983-5	Sequence 5, Appli
17	651	100.0	3030	16	US-10-369-983-3	Sequence 3, Appli
18	651	100.0	3060	16	US-10-369-983-11	Sequence 11, Appl
19	651	100.0	3104	16	US-10-369-983-10	Sequence 10, Appl
20	651	100.0	3474	16	US-10-369-983-9	Sequence 9, Appli
21	646	99.2	447	15	US-10-193-002-4	Sequence 4, Appli
22	646	99.2	447	15	US-10-084-843-4	Sequence 4, Appli
23	646	99.2	447	15	US-10-098-732A-9	Sequence 9, Appli
24	646	99.2	1872	15	US-10-193-002-17	Sequence 17, Appl
25	646	99.2	1872	15	US-10-084-843-17	Sequence 17, Appl
26	646	99.2	1872	15	US-10-098-732A-1	Sequence 1, Appli
27	608	93.4	675	9	US-09-759-143-822	Sequence 822, App
28	608	93.4	675	9	US-09-780-669-822	Sequence 822, App
29	608	93.4	675	9	US-09-822-827-822	Sequence 822, App
30	608	93.4	675	9	US-09-895-793-822	Sequence 822, App
31	608	93.4	675	9	US-09-895-814-822	Sequence 822, App
32	608	93.4	675	14	US-10-012-896-822	Sequence 822, App
33	608	93.4	675	15	US-10-144-678A-822	Sequence 822, App
34	608	93.4	675	15	US-10-294-025-822	Sequence 822, App
35	608	93.4	822	9	US-09-736-457-1862	Sequence 1862, Ap
36	608	93.4	822	9	US-09-902-941-1862	Sequence 1862, Ap
37	608	93.4	822	9	US-09-849-626-1862	Sequence 1862, Ap
38	608	93.4	822	13	US-10-283-017-1862	Sequence 1862, Ap
39	608	93.4	822	15	US-10-017-754-1862	Sequence 1862, Ap
40	608	93.4	822	15	US-10-113-872-1862	Sequence 1862, Ap
41	608	93.4	894	13	US-09-878-722-236	Sequence 236, App
42	608	93.4	894	13	US-09-904-456-236	Sequence 236, App
43	608	93.4	900	9	US-09-735-705-353	Sequence 353, App
44	608	93.4	900	9	US-09-850-716A-353	Sequence 353, App
45	608	93.4	900	9	US-09-897-778-353	Sequence 353, App

ALIGNMENTS

RESULT 1
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556

; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 702
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
;   OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
;   OTHER INFORMATION: frame 1
;   NAME/KEY: CDS
;   LOCATION: (1)..(693)
;   OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
;   OTHER INFORMATION: reading frame 1
;   NAME/KEY: CDS
;   LOCATION: (2)..(700)
;   OTHER INFORMATION: reading frame 2
;   NAME/KEY: CDS
;   LOCATION: (3)..(701)
;   OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:

Pred. No.:	3.4e-69	Length:	702
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-23 (1-128) x US-09-287-849-27 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 2
US-10-359-460-27
; Sequence 27, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 702
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
;   OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
;   OTHER INFORMATION: frame 1
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(693)
;   OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
;   OTHER INFORMATION: reading frame 1
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(700)
;   OTHER INFORMATION: reading frame 2
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(701)
;   OTHER INFORMATION: reading frame 3
US-10-359-460-27

Alignment Scores:

Pred. No.:	3.4e-69	Length:	702
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-23 (1-128) x US-10-359-460-27 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324
```

```
Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 3
US-10-098-732A-3
; Sequence 3, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Ra35 mature
US-10-098-732A-3

Alignment Scores:
Pred. No.:              5.25e-69        Length:         1002
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-098-732A-3 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAG 981

RESULT 4
US-10-098-732A-5
; Sequence 5, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Ra35FLMutSA
US-10-098-732A-5

Alignment Scores:
Pred. No.:              5.25e-69        Length:         1002
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-098-732A-5 (1-1002)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAG 981

RESULT 5
US-09-712-363-15
; Sequence 15, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
```

```
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE: 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-15

Alignment Scores:
Pred. No.:                5.67e-69        Length:          1068
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-09-684-215B-23 (1-128) x US-09-712-363-15 (1-1068)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAG 1053

RESULT 6
US-10-369-983-1
; Sequence 1, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT: Guderian, Jeff
;  APPLICANT: Reed, Steven
;  APPLICANT: Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2181
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:mutated
;   OTHER INFORMATION: MTB32-MTB39F fusion protein (MTB32MutSA)
US-10-369-983-1

Alignment Scores:
Pred. No.:                1.35e-68        Length:          2181
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       16              Gaps:            0

US-09-684-215B-23 (1-128) x US-10-369-983-1 (1-2181)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 657

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 717

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 777

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 837

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 897

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 957

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||||
Db        958 ACAGGGAACGTGACATTGGCCGAG 981

RESULT 7
US-10-098-732A-17
; Sequence 17, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
```

; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 2190
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB72FMutSA
; OTHER INFORMATION: (Ra12-TbH9-Ra35MutSA)
US-10-098-732A-17

Alignment Scores:

Pred. No.:	1.36e-68	Length:	2190
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-23 (1-128) x US-10-098-732A-17 (1-2190)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 141

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 201

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 261

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 321

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 381

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        382 ACAGGGAACGTGACATTGGCCGAG 405
```

RESULT 8
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2287
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
; OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
; OTHER INFORMATION: fusion)
; NAME/KEY: modified_base
; LOCATION: (30)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: modified_base
; LOCATION: (33)
; OTHER INFORMATION: n = g, a, c or t
; NAME/KEY: CDS
; LOCATION: (42)..(2231)
; NAME/KEY: modified_base
; LOCATION: (2270)
; OTHER INFORMATION: n = g, a, c or t
US-09-287-849-1

Alignment Scores:

Pred. No.:	1.43e-68	Length:	2287
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-23 (1-128) x US-09-287-849-1 (1-2287)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446
```

RESULT 9
US-10-359-460-1
; Sequence 1, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark

```
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE: 014058-009020US
;  CURRENT APPLICATION NUMBER: US/10/359,460
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US/09/287,849
;  PRIOR FILING DATE: 1999-04-07
;  PRIOR APPLICATION NUMBER: US 08/818,112
;  PRIOR FILING DATE: 1997-03-13
;  PRIOR APPLICATION NUMBER: US 08/942,578
;  PRIOR FILING DATE: 1997-10-01
;  PRIOR APPLICATION NUMBER: US 09/025,197
;  PRIOR FILING DATE: 1998-02-18
;  PRIOR APPLICATION NUMBER: US 09/056,556
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: US 09/223,040
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
;   OTHER INFORMATION: fusion)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-10-359-460-1

Alignment Scores:
Pred. No.:                1.43e-68        Length:          2287
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-23 (1-128) x US-10-359-460-1 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 10
US-10-098-732A-15
; Sequence 15, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein MTB72F (Ra12-TbH9-Ra35 or MTB32-MTB39
;   OTHER INFORMATION: fusion)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   OTHER INFORMATION: MTB72F
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1)..(2287)
;   OTHER INFORMATION: n = g, a, c or t
US-10-098-732A-15

Alignment Scores:
Pred. No.:                1.43e-68        Length:          2287
Score:                    651.00          Matches:         128
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-23 (1-128) x US-10-098-732A-15 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302
```

```
Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
               |||||||||||||||||||||||
Db         423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 11
US-10-359-459-1
; Sequence 1, Application US/10359459
; Publication No. US20040013677A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE: 014058-009010US
;  CURRENT APPLICATION NUMBER: US/10/359,459
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US/09/223,040
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ra12-TbH9-Ra35
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-10-359-459-1

Alignment Scores:
Pred. No.:              1.43e-68        Length:         2287
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-359-459-1 (1-2287)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
               |||||||||||||||||||||||
Db         423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 12
US-10-369-983-8
; Sequence 8, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 2451
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB81F (MTB72F-DPV)
US-10-369-983-8

Alignment Scores:
Pred. No.:              1.56e-68        Length:         2451
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-369-983-8 (1-2451)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 13
US-10-369-983-7
; Sequence 7, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: MTB83F (MTB72F-MTI)
US-10-369-983-7

Alignment Scores:

Pred. No.:	1.59e-68	Length:	2487
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	16	Gaps:	0

US-09-684-215B-23 (1-128) x US-10-369-983-7 (1-2487)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 14
US-10-369-983-6
; Sequence 6, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: MTB89F (MTB72F-Erd14)
US-10-369-983-6

Alignment Scores:

Pred. No.:	1.7e-68	Length:	2637
Score:	651.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	16	Gaps:	0

US-09-684-215B-23 (1-128) x US-10-369-983-6 (1-2637)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408
```

RESULT 15
US-10-098-732A-64
; Sequence 64, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey

```
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTB72F-MAPS
;   OTHER INFORMATION: (r95f) fusion construct, TB MTB72F (Ra12-TbH9-Ra35)
;   OTHER INFORMATION: linked to Leishmania thiol-specific antioxidant
;   OTHER INFORMATION: (TSA or MAPS)
US-10-098-732A-64

Alignment Scores:
Pred. No.:                1.84e-68        Length:         2808
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       15              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-098-732A-64 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 16
US-10-369-983-5
; Sequence 5, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 2808
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: R95F (MTB72F-MAPS)
US-10-369-983-5

Alignment Scores:
Pred. No.:                1.84e-68        Length:         2808
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       16              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-369-983-5 (1-2808)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 17
US-10-369-983-3
; Sequence 3, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 3030
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTB-102F fusion
```

```
;   OTHER INFORMATION: protein
US-10-369-983-3

Alignment Scores:
Pred. No.:                2.02e-68        Length:         3030
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       16              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-369-983-3 (1-3030)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 654

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 714

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 774

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 834

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 894

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 954

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        955 ACAGGGAACGTGACATTGGCCGAG 978

RESULT 18
US-10-369-983-11
; Sequence 11, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 3060
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB103F (MTB72F-85b)
US-10-369-983-11

Alignment Scores:
Pred. No.:                2.04e-68        Length:         3060
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       16              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-369-983-11 (1-3060)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 19
US-10-369-983-10
; Sequence 10, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 3104
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB102tm2F (MTB102FTM, MTB72F-hTCC#1)
US-10-369-983-10

Alignment Scores:
Pred. No.:                2.08e-68        Length:         3104
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       16              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-369-983-10 (1-3104)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
                |||||||||||||||||||||||
Db         385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 20
US-10-369-983-9
; Sequence 9, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Guderian, Jeff
;  APPLICANT:  Reed, Steven
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
;  FILE REFERENCE: 014058-009081US
;  CURRENT APPLICATION NUMBER: US/10/369,983
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US 60/357,351
;  PRIOR FILING DATE: 2002-02-15
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 3474
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:fusion protein
;   OTHER INFORMATION: MTB114F (MTB72F-mTCC#2)
US-10-369-983-9

Alignment Scores:
Pred. No.:              2.39e-68        Length:         3474
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     16              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-369-983-9 (1-3474)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
                |||||||||||||||||||||||
Db         385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 21
US-10-193-002-4
; Sequence 4, Application US/10193002
; Publication No. US20030135026A1
;  GENERAL INFORMATION:
;       APPLICANT: Reed, Steven G.
;                  Skeiky, Yasir A.W.
;                  Dillon, Davin C.
;                  Campos-Neto, Antonia
;                  Houghton, Raymond
;                  Vedvick, Thomas S.
;                  Twardzik, Daniel R.
;                  Lodes, Michael J.
;                  Hendrickson, Ronald C.
;       TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                           TUBERCULOSIS
;       NUMBER OF SEQUENCES: 350
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SEED and BERRY LLP
;            STREET: 6300 Columbia Center, 701 Fifth Avenue
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/193,002
;            FILING DATE: 10-Jul-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/072,596
;            FILING DATE: 05-MAY-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Maki, David J.
;            REGISTRATION NUMBER: 31,392
;            REFERENCE/DOCKET NUMBER: 210121.417C9
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 447 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-193-002-4

Alignment Scores:
Pred. No.:              7.87e-69        Length:         447
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     15              Gaps:           0
```

```
US-09-684-215B-23 (1-128) x US-10-193-002-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 22
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonio
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;                            AND DIAGNOSIS OF TUBERCULOSIS
;        NUMBER OF SEQUENCES: 355
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/084,843
;             FILING DATE: 25-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,967
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.411C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 447 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-084-843-4

Alignment Scores:
Pred. No.:                7.87e-69        Length:          447
Score:                    646.00          Matches:         127
Percent Similarity:       99.22%          Conservative:    0
Best Local Similarity:    99.22%          Mismatches:      1
Query Match:              99.23%          Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-23 (1-128) x US-10-084-843-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 23
US-10-098-732A-9
; Sequence 9, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
;  TITLE OF INVENTION:  Leishmania Antigen
;  FILE REFERENCE: 014058-012010US
;  CURRENT APPLICATION NUMBER: US/10/098,732A
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: US 60/275,837
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 447
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:MTBRa12 (MTB32A
;   OTHER INFORMATION: C-terminus)
```

```
US-10-098-732A-9

Alignment Scores:
Pred. No.:              7.87e-69        Length:         447
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-098-732A-9 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 24
US-10-193-002-17
; Sequence 17, Application US/10193002
; Publication No. US20030135026A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonia
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                            TUBERCULOSIS
;        NUMBER OF SEQUENCES: 350
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/193,002
;             FILING DATE: 10-Jul-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,596
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.417C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 17:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1872 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-193-002-17

Alignment Scores:
Pred. No.:              4.5e-68         Length:         1872
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-23 (1-128) x US-10-193-002-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 25
US-10-084-843-17
; Sequence 17, Application US/10084843
; Publication No. US20030143243A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonio
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;                            AND DIAGNOSIS OF TUBERCULOSIS
```

; NUMBER OF SEQUENCES: 355
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/084,843
; FILING DATE: 25-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,967
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-084-843-17

Alignment Scores:

Pred. No.:	4.5e-68	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-23 (1-128) x US-10-084-843-17 (1-1872)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

RESULT 26
US-10-098-732A-1
; Sequence 1, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: MTB32A (Ra35FL)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(1872)
; OTHER INFORMATION: n = g, a, c or t
US-10-098-732A-1

Alignment Scores:

Pred. No.:	4.5e-68	Length:	1872
Score:	646.00	Matches:	127
Percent Similarity:	99.22%	Conservative:	0
Best Local Similarity:	99.22%	Mismatches:	1
Query Match:	99.23%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-23 (1-128) x US-10-098-732A-1 (1-1872)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

RESULT 27
US-09-759-143-822
; Sequence 822, Application US/09759143
; Patent No. US20020022248A1

```
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C23
;  CURRENT APPLICATION NUMBER: US/09/759,143
;  CURRENT FILING DATE:  2001-01-12
;  NUMBER OF SEQ ID NOS:  934
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-759-143-822

Alignment Scores:
Pred. No.:                4.97e-64        Length:          675
Score:                    608.00          Matches:         122
Percent Similarity:       96.09%          Conservative:    1
Best Local Similarity:    95.31%          Mismatches:      1
Query Match:              93.39%          Indels:          4
DB:                       9               Gaps:            1

US-09-684-215B-23 (1-128) x US-09-759-143-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 28
US-09-780-669-822
; Sequence 822, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.427C24
;  CURRENT APPLICATION NUMBER: US/09/780,669
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS:  943
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-780-669-822

Alignment Scores:
Pred. No.:                4.97e-64        Length:          675
Score:                    608.00          Matches:         122
Percent Similarity:       96.09%          Conservative:    1
Best Local Similarity:    95.31%          Mismatches:      1
Query Match:              93.39%          Indels:          4
DB:                       9               Gaps:            1

US-09-684-215B-23 (1-128) x US-09-780-669-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 29
```

```
US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C1
;  CURRENT APPLICATION NUMBER: US/09/822,827
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-822-827-822

Alignment Scores:
Pred. No.:              4.97e-64        Length:         675
Score:                  608.00          Matches:        122
Percent Similarity:     96.09%          Conservative:   1
Best Local Similarity:  95.31%          Mismatches:     1
Query Match:            93.39%          Indels:         4
DB:                     9               Gaps:           1

US-09-684-215B-23 (1-128) x US-09-822-827-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::             ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 30
US-09-895-793-822
; Sequence 822, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.534C2
;  CURRENT APPLICATION NUMBER: US/09/895,793
;  CURRENT FILING DATE:  2001-06-29
;  NUMBER OF SEQ ID NOS: 982
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-895-793-822

Alignment Scores:
Pred. No.:              4.97e-64        Length:         675
Score:                  608.00          Matches:        122
Percent Similarity:     96.09%          Conservative:   1
Best Local Similarity:  95.31%          Mismatches:     1
Query Match:            93.39%          Indels:         4
DB:                     9               Gaps:           1

US-09-684-215B-23 (1-128) x US-09-895-793-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::             ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393


Search completed: April 30, 2004, 05:10:29
Job time : 474.379 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:25:27 ; Search time 2507.02 Seconds
(without alignments)
1524.660 Million cell updates/sec

Title: US-09-684-215B-23
Perfect score: 651
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....SVTWQTKSGGTRTGNVTLAE 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061305_13200/app_query.fasta_1.1180
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_5167_@runat_29042004_061305_13200 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		145.5	22.4	289	29	U82114	U82114 U82114 orde
2	c	126.5	19.4	1438	28	BH770798	BH770798 LLMGtag54
3	c	122	18.7	959	28	BZ549048	BZ549048 pacs1-60_
4	c	120.5	18.5	726	14	CB679186	CB679186 OSJNBf02F
5		117.5	18.0	590	14	CB925602	CB925602 ABA1_22_F
6	c	115.5	17.7	807	14	CA228900	CA228900 SCQGFL305
7		114	17.5	603	28	AZ934428	AZ934428 BJ__Ba000
8	c	113	17.4	758	13	BQ514888	BQ514888 EST622303
9		108	16.6	982	14	CD080132	CD080132 MA3-9999U
10	c	106.5	16.4	1033	28	BZ561390	BZ561390 pacs2-164
11	c	103.5	15.9	947	28	BZ549047	BZ549047 pacs1-60_
12		103	15.8	726	28	AQ989479	AQ989479 Rfc00025_
13		102	15.7	719	28	AZ933900	AZ933900 BJ__Ba000
14		102	15.7	765	9	AJ558965	AJ558965 AJ558965
15	c	100.5	15.4	859	14	CA480614	CA480614 AGENCOURT
16		99	15.2	423	9	AI597611	AI597611 tn15f02.x
17	c	97.5	15.0	499	12	BI350520	BI350520 fr32b05.y
18	c	97.5	15.0	870	14	CA473840	CA473840 AGENCOURT
19		97	14.9	444	13	CA148171	CA148171 SCEZRZ101
20		97	14.9	543	12	BI721127	BI721127 1031054B1
21		96	14.7	546	10	AW285510	AW285510 LG1_241_E
22		96	14.7	551	10	AW285527	AW285527 LG1_241_G
23		95.5	14.7	677	12	BI378928	BI378928 BFLG1_000
24	c	95	14.6	909	14	CA975497	CA975497 AGENCOURT
25	c	94.5	14.5	600	14	CF787130	CF787130 854999 MA
26	c	94.5	14.5	647	13	BX674871	BX674871 BX674871
27	c	94.5	14.5	728	12	BJ285991	BJ285991 BJ285991
28		93.5	14.4	600	28	CC345457	CC345457 OGQAF70TH
29	c	93.5	14.4	907	14	CA471201	CA471201 AGENCOURT
30	c	93.5	14.4	913	14	CF265501	CF265501 AGENCOURT
31	c	93	14.3	951	14	CA474469	CA474469 AGENCOURT
32	c	92.5	14.2	709	29	CG895905	CG895905 pastbac00
33	c	92.5	14.2	860	14	CA471291	CA471291 AGENCOURT
34	c	92.5	14.2	949	14	CF243195	CF243195 AGENCOURT
35		92.5	14.2	1394	28	BZ576076	BZ576076 msh2_4761
36		92	14.1	766	14	CF997044	CF997044 AGENCOURT
37	c	92	14.1	790	14	CK127722	CK127722 AGENCOURT
38	c	92	14.1	894	28	BZ578796	BZ578796 msh2_6009
39	c	92	14.1	899	14	CA471516	CA471516 AGENCOURT
40		91.5	14.1	452	12	BM874057	BM874057 laa07a02.
41	c	91.5	14.1	513	12	BM036648	BM036648 fu80g01.y
42	c	91.5	14.1	555	13	BQ479512	BQ479512 faa68g09.
43	c	91.5	14.1	570	14	CB364397	CB364397 ZF001-P00
44	c	91.5	14.1	613	10	AW595047	AW595047 fk28e03.y
45	c	91.5	14.1	631	14	CF512107	CF512107 CAbud0004

ALIGNMENTS

RESULT 1
U82114
LOCUS U82114 289 bp DNA linear GSS 16-FEB-2001
DEFINITION U82114 ordered cosmid library Mycobacterium leprae genomic clone cosmid L-373; contig 64, genomic survey sequence.
ACCESSION U82114
VERSION U82114.1 GI:3647212
KEYWORDS GSS.
SOURCE Mycobacterium leprae
ORGANISM Mycobacterium leprae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
REFERENCE 1 (bases 1 to 289)

```
  AUTHORS   Silbaq,F.S., Cho,S.N., Cole,S.T. and Brennan,P.J.
  TITLE     Characterization of a 34-kilodalton protein of Mycobacterium leprae
            that is isologous to the immunodominant 34-kilodalton antigen of
            Mycobacterium paratuberculosis
  JOURNAL   Infect. Immun. 66 (11), 5576-5579 (1998)
  MEDLINE   99003183
   PUBMED   9784577
COMMENT     Contact: Silbaq FS
            Microbiology
            Colorado State University
            Fort Collins, CO 80523, USA
            Eiglmeier,K., Honore,N., Woods,S.A., Caudron,B. and Cole,S.T. Use
            of an ordered cosmid library to deduce the genomic organization of
            Mycobacterium leprae. Mol. Microbiol. 7 (2), 197-206 (1993)
            Class: unknown.
FEATURES             Location/Qualifiers
     source          1. .289
                     /organism="Mycobacterium leprae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1769"
                     /clone="cosmid L-373; contig 64"
                     /clone_lib="ordered cosmid library"
ORIGIN

Alignment Scores:
Pred. No.:                6.63e-05        Length:         289
Score:                    145.50          Matches:        35
Percent Similarity:       58.82%          Conservative:   15
Best Local Similarity:    41.18%          Mismatches:     34
Query Match:              22.35%          Indels:         1
DB:                       29              Gaps:           1

US-09-684-215B-23 (1-128) x U82114 (1-289)

Qy          45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn---GlyAlaArgValGlnArg 63
               |||   ||||||:::  |||      :::  |||      ||||||:::|||
Db           7 GCGTCGCTGGGTGTACAGGTGGCCACCGACAAAGGCACCCCGGGCGCCAAAGTTATGGAC 66

Qy          64 ValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 83
               ||||||              ||||||:::     :::      |||   |||:::|||
Db          67 GTCGTGGCCGGTGGTGCGGCCGCGAATGCCGCCGTTCCCAAGGGGGTCGTTCTCACTAAG 126

Qy          84 ValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHis 103
               ||||||        |||:::||||||   |||:::      |||:::
Db         127 GTCGACGACCGCCTGATCAGTAGCGCCGACGCGTTGGTCGCTGCCGTGCGGTCCAAGGCA 186

Qy         104 ProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsn 123
               |||||||||   :::|||:::|||:::|||   :::||||||   :::||||||
Db         187 CCCGGTGACAAAGTGTCGCTGACCTATCAGGATCAGTCTGGTAGCAGTCGCACGGTTCAG 246

Qy         124 ValThrLeuAlaGlu 128
               |||||||||   :::
Db         247 GTCACACTCGGCAAG 261

RESULT 2
BH770798/c
LOCUS       BH770798               1438 bp    DNA     linear   GSS 01-MAY-2002
DEFINITION  LLMGtag541 MG1363 Random Sequence Tag Library Lactococcus lactis
            subsp. cremoris genomic, genomic survey sequence.
ACCESSION   BH770798
VERSION     BH770798.1  GI:20373755
KEYWORDS    GSS.
SOURCE      Lactococcus lactis subsp. cremoris
  ORGANISM  Lactococcus lactis subsp. cremoris
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Lactococcus.
REFERENCE   1  (bases 1 to 1438)
  AUTHORS   Bolotin,A., Ehrlich,S.D. and Sorokin,A.
  TITLE     Studies of genomes of dairy bacteria Lactococcus lactis
  JOURNAL   Sci. Aliments (2002) In press
COMMENT     Contact: Sorokin A
            Genetique Microbienne
            INRA
            CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
            Tel: 33 1 34 65 25 16
            Fax: 33 1 34 65 25 21
            Email: sorokine@jouy.inra.fr
            best homologue in strain IL1403 is htrA (95%)
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 1408.
FEATURES             Location/Qualifiers
     source          1. .1438
                     /organism="Lactococcus lactis subsp. cremoris"
                     /mol_type="genomic DNA"
                     /strain="MG1363"
                     /sub_species="cremoris"
                     /db_xref="taxon:1359"
                     /clone_lib="MG1363 Random Sequence Tag Library"
                     /note="Vector: pSGMU2; Site_1: SmaI; Library of
                     chromosomal fragments of L.lactis strain MG1363 was
                     prepared by partial AluI digestion or by sonication."
ORIGIN

Alignment Scores:
Pred. No.:                0.0521          Length:         1438
Score:                    126.50          Matches:        40
Percent Similarity:       45.45%          Conservative:   20
Best Local Similarity:    30.30%          Mismatches:     49
Query Match:              19.43%          Indels:         23
DB:                       28              Gaps:           5

US-09-684-215B-23 (1-128) x BH770798 (1-1438)

Qy          13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
               |||   |||||||||||||||           :::   |||     ::::::   :::
Db         479 GGTTTAGGTTTTGCTATCCCATCTAATGATGTGGTAAACATCATTAATAAACTTGAAACT 420

Qy          33 GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal 52
               |||                 |||          |||   ||||||:::   :::|||
Db         419 GATGGT--------------AAGATTTCACGTCCTGCC---TTAGGTATTCGTATGGTT 378

Qy          53 Asp-----------------AsnAsnGlyAsn-------------------------- 57
               |||                 |||:::   :::
Db         377 GACCTATCTCAATTATCAACAAATGATAGTTCTCAACTGAAATTACCTAGCAGCGTAACT 318

Qy          58 ---GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSer 76
                  |||     |||       |||                 ||||||:::   |||:::
Db         317 GGTGGAGTGGTTGTCTACTCTGTTCAAGCGGGTCTTTCTGCTGCCACAGCTGGTCTGAAA 258

Qy          77 ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAla 96
               |||||||||||||||   |||          :::   |||:::|||   :::
Db         257 GCTGGCGATGTGATAACGAAGGTGGGAGATACCGCCGTTACTTCATCAACAGACTTACAA 198

Qy          97 AspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSer 116
               ||||||      |||:::      |||   :::   ||||||:::       :::
Db         197 AGTGCTCTTTACTCACACAATATTAATGATACTGTGAAAGTCACCTACTAC---CGTGAT 141

Qy         117 GlyGlyThrArgThrGlyAsnValThrLeuAlaGlu 128
               |||   :::   |||   ||||||   |||::::::
Db         140 GGTAAATCAGCCACAGCAAATGTCAAACTCTCTAAA 105

RESULT 3
BZ549048/c
LOCUS       BZ549048                959 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_1644.s2 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_1644, genomic survey sequence.
ACCESSION   BZ549048
VERSION     BZ549048.1  GI:27152629
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 959)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
```

```
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .959
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="1-60"
                     /db_xref="taxon:287"
                     /clone="pacs1-60_1644"
                     /clone_lib="pacs1-60"
                     /note="clinical isolate 1-60 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:                0.0809          Length:         959
Score:                    122.00          Matches:        40
Percent Similarity:       47.29%          Conservative:   21
Best Local Similarity:    31.01%          Mismatches:     44
Query Match:              18.74%          Indels:         24
DB:                       28              Gaps:           5

US-09-684-215B-23 (1-128) x BZ549048 (1-959)

Qy          4 SerAspAsnPheGlnLeuSerGlnGly------GlyGlnGlyPheAlaIleProIleGly 21
              |||     |||      |||   ||       |||      |||||||||||||||
Db        378 TCGCAGATCTTCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGAT 319

Qy         22 GlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle 41
                ||:::   :::|||   |||::::::     |||                  |||
Db        318 GTCGCGCTGAACGTCGCCGACCAGTTGAAGAAAGCCGGC---------AAGGTCAGTCGC 268

Qy         42 GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsn------------------ 55
              |||         :::||||||:::   :::   :::   |||
Db        267 GGC---------TGGCTGGGCGTGGTGATCCAGGAAGTGAACAAGGATCTCGCCGAGTCC 217

Qy         56 --------------GlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70
                           :::|||||   |||   ::::::|||         ||||||
Db        216 TACGGCCTCGACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCG 157

Qy         71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 90
              |||      |||:::     |||||||||||   :::::::::|||      ||||||
Db        156 GCCAAGGGTGGCCTGCAGGTGGGCGATGTGATCATAAGCTTGAACGGCCAGTCGATCAAC 97

Qy         91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
                :::     :::     :::                |||||||||   |||::::::
Db         96 GAGTCCGCCGACCTGCCGCACCTGGTGGGCAACATGAAGCCGGGCGACAAGATCAACCTG 37

Qy        111 ThrTrpGlnThrLysSerGlyGlyThr 119
                          ||||||:::
Db         36 --------------GACGGGGGATCC 25

RESULT 4
CB679186/c
LOCUS       CB679186                 726 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEf02F19.r OSJNEf Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEf02F19 3', mRNA sequence.
ACCESSION   CB679186
VERSION     CB679186.1  GI:29682911
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 02  row: F  column: 19
            Seq primer: gga aac agc tat gac cat g.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEf02F19"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEf"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; Uninfected Control"
ORIGIN

Alignment Scores:
Pred. No.:                0.0761          Length:         726
Score:                    120.50          Matches:        41
Percent Similarity:       44.03%          Conservative:   18
Best Local Similarity:    30.60%          Mismatches:     50
Query Match:              18.51%          Indels:         25
DB:                       14              Gaps:           4

US-09-684-215B-23 (1-128) x CB679186 (1-726)

Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla------ 27
              :::      |||   |||||||||||||||           :::   ||||||
Db        579 ACATCTGCTGGTGTTGGTTTTGCCATCCCGTCATCAACTGTTCTGAAAATAGCTCCTCAG 520

Qy         28 ------------GlyGlnIleArgSerGlyGlyGlySerProThrValHisIleGlyPro 43
                         |||:::::|||      |||  :::                  |||
Db        519 TTAATTCAGTTTGGAAAAGTTCGCCGTGCTGGCTTGAATGTGGAGTTTGCTCCAGATCCA 460

Qy         44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63
              ||:::                     |||          |||:::   :::   :::
Db        459 ATCGCATAT--------------CAGCTTAATGTTCGCACTGGATCTATAATATTGCAG 415

Qy         64 ValValGlySerAlaProAlaAlaSerLeuGly-------------------------- 74
              |||   |||   :::   ||||||      |||
Db        414 GTTCCTGGGGGCAGTGCTGCAGCCAAAGCTGGTCTTGTTCCTACCAGTAGGGGTTTTGCT 355

Qy         75 ------IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
                   |||      ||||||||||||   ||||||||||   ||||||
Db        354 GGTACCATTGTTCTTGGTGATGTCATTGTTGCCGTGGACGGTAAACCTATCAAAGGCAAA 295

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112
              :::   ::::::      |||:::   :::      ||||||   :::|||:::|||
Db        294 TCTGATCTATCGAGGGTTCTGGATGCTTATGGCGTTGGAGACAAGGTGAGCTTGACAATC 235

Qy        113 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 126
              |||      |||      :::      |||||||||
Db        234 CAAAGA---------GGCGCTGAATCTCTGGAGGTAACCTTG 202
```

```
RESULT 5
CB925602
LOCUS       CB925602                590 bp    mRNA    linear   EST 28-APR-2003
DEFINITION  ABA1_22_F06.b1_A012 Abscisic acid-treated seedlings Sorghum bicolor
            cDNA clone ABA1_22_F06_A012 3', mRNA sequence.
ACCESSION   CB925602
VERSION     CB925602.1  GI:30161873
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S.,
            Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Buchanan,C.D.,
            Eastman,A. and Pratt,L.H.
  TITLE     An EST database from Sorghum: ABA1-treated seedlings
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: ABA1_22_F06.g1_A012
            Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
            the Human Genome Center, University of Tokyo Institute of Medical
            Science; plant material and RNA prepared at Texas A & M University;
            sequencing done in the Laboratory for Genomics and Bioinformatics,
            University of Georgia. Sequence ends have been trimmed to exclude
            vector and regions below Phred quality 16. Three-prime sequences
            are presented as their reverse complement and have been trimmed to
            exclude polyA.
            Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="IS3620C"
                     /db_xref="taxon:4558"
                     /clone="ABA1_22_F06_A012"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Abscisic acid-treated seedlings"
                     /note="Vector: pME18S-FL3; Site_1: XhoI; Site_2: XhoI; The
                     library was prepared from polyA+ RNA from seedlings grown
                     in hydroponic culture. After 12 days, medium was
                     supplemented with 1 mM abscisic acid (ABA), while leaves
                     were misted with a solution of 1 mM ABA. Roots and leaves
                     were harvested after 3, 6, 12, and 24 hr and material from
                     all time points was combined prior to RNA isolation.
                     Double-stranded cDNA was cloned unidirectionally into
                     different DraIII sites of the pME18S-FL3 vector (5-prime
                     DraIII site is CACTGTGTG, 3-prime DraIII site is
                     CACCATGTG). XhoI excises the cDNA insert."
ORIGIN

Alignment Scores:
Pred. No.:                0.112          Length:         590
Score:                    117.50         Matches:        37
Percent Similarity:       44.19%         Conservative:   20
Best Local Similarity:    28.68%         Mismatches:     43
Query Match:              18.05%         Indels:         29
DB:                       14             Gaps:           4

US-09-684-215B-23 (1-128) x CB925602 (1-590)

Qy          15 GlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla------------------Gly 28
               ||||||||||||||||           :::   ||||||                 |||
Db           6 GGCTTTGCTATTCCATCATCAACTGTACTTAAAATCGCTCCTCAGTTAATTCAGTTCGGA 65

Qy          29 GlnIleArgSerGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGly 48
               ::::::|||       |||   :::                 |||   |||:::
Db          66 AAAGTTCGTCGTGCTGGCTTGAATGTGGACTTCGCTCCGGATCCAATTGCATAT------ 119

Qy          49 LeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAla 68
                        |||       |||||||||   :::   :::|||   |||   :::
Db         120 ---------CAGCTTAATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGT 170

Qy          69 ProAlaAlaSerLeuGlyIleSer--------------------------------Thr 77
                  ||||||       |||::::::
Db         171 GCTGCAGCCAAAGCAGGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTG 230

Qy          78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
               |||||:::|||   |||||||||||||   |||:::         :::   :::
Db         231 GGTGATATCATCGTTGCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGG 290

Qy          98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGly 117
                  |||:::   :::      ||||||   :::::::::|||
Db         291 GTTCTGGATGACTATGGCGTCGGAGATCAGGTGACCTTGACA---------------- 332

Qy         118 GlyThrArgThrGlyAsnValThrLeu 126
                  |||   |||:::   ||||||
Db         333 ---ATCCGGCGAGGCTCAGAAACCCTT 356

RESULT 6
CA228900/c
LOCUS       CA228900                807 bp    mRNA    linear   EST 25-SEP-2003
DEFINITION  SCQGFL3058G06.g Saccharum officinarum FL3 Saccharum officinarum
            cDNA clone SCQGFL3058G06 5', mRNA sequence.
ACCESSION   CA228900
VERSION     CA228900.1  GI:35290878
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1  (bases 1 to 807)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 058  row: G  column: 06
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .807
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCQGFL3058G06"
                     /lab_host="DH10B"
                     /clone_lib="Saccharum officinarum FL3"
                     /note="Organ: Base of developing inflorescence (5cm-long);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Base of
                     developing inflorescence (5cm-long)]. cDNA was prepared
                     from polyA+ mRNA using SuperScript Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
```

```
ORIGIN

Alignment Scores:
Pred. No.:              0.278          Length:         807
Score:                  115.50         Matches:        38
Percent Similarity:     43.28%         Conservative:   20
Best Local Similarity:  28.36%         Mismatches:     47
Query Match:            17.74%         Indels:         29
DB:                     14             Gaps:           4

US-09-684-215B-23 (1-128) x CA228900 (1-807)

Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAla------ 27
              :::      |||   |||||||||||          :::   ||||||
Db        569 ACATCTGCTGGTGTTGGCTTCGCTATCTCATCATCAACTGTGCTTAAAATCGCTCCTCAG 510

Qy         28 ------------GlyGlnIleArgSerGlyGlyGlySerProThrValHisIleGlyPro 43
                        |||::::::|||      |||   :::                  |||
Db        509 TTAATTCAGTTTGGAAAAGTTCGTCGTGCTGGCTTGAATGTGGACTTTGCTCCAGATCCA 450

Qy         44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63
                 |||:::                  |||      |||||||||   :::   :::
Db        449 ATTGCGTAT---------------CAGCTTAATGTTCGCAACGGAGCTCTTATACTTAAG 405

Qy         64 ValValGlySerAlaProAlaAlaSerLeuGlyIleSer--------------------- 76
              |||   |||   :::   ||||||      |||::::::
Db        404 GTACCTGGGGGCAGTGCTGCAGCCAAGGCAGGCCTTGCTCCGACCGGCAGGGGTTTCGCT 345

Qy         77 ------------ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
                        |||||||||||   ||||||||||   |||:::
Db        344 GGTAACATTGTTCTTGGTGATGTCATCGTTGCAGTGGATGGCAAACCTGTTAAGGGCAAA 285

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112
              :::   :::        |||:::   :::      ||||||   :::::::::|||
Db        284 TCTGACCTGCTGAGGGTTCTGGATGACTATGGCGTCGGGGATCAGGTGACCTTGACA--- 228

Qy        113 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeu 126
                          |||   |||:::   ||||||
Db        227 -----------------ATCCGGCGAGGCTCTGAAACCCTC 204

RESULT 7
AZ934428
LOCUS       AZ934428                603 bp    DNA     linear   GSS 24-APR-2001
DEFINITION  BJ__Ba0002I08r B. japonicum BAC library Bradyrhizobium japonicum
            genomic, genomic survey sequence.
ACCESSION   AZ934428
VERSION     AZ934428.1  GI:13776488
KEYWORDS    GSS.
SOURCE      Bradyrhizobium japonicum
  ORGANISM  Bradyrhizobium japonicum
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1  (bases 1 to 603)
  AUTHORS   Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
            Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
  TITLE     A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
            genome
  JOURNAL   Genome Res. 11 (8), 1434-1440 (2001)
  MEDLINE   21376150
   PUBMED   11483585
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Class: BAC ends
            High quality sequence stop: 553.
FEATURES             Location/Qualifiers
     source          1. .603
                     /organism="Bradyrhizobium japonicum"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:375"
                     /lab_host="E. coli"
                     /clone_lib="B. japonicum BAC library"
                     /note="Vector: pIndigo536; Site_1: HindIII"
ORIGIN

Alignment Scores:
Pred. No.:              0.257          Length:         603
Score:                  114.00         Matches:        36
Percent Similarity:     43.44%         Conservative:   17
Best Local Similarity:  29.51%         Mismatches:     57
Query Match:            17.51%         Indels:         12
DB:                     28             Gaps:           4

US-09-684-215B-23 (1-128) x AZ934428 (1-603)

Qy         13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
              |||      |||:::||||||                    :::      |||::::::
Db        245 GGCATCGCGTTCTCGATCCCGGCCAACACCGTGAAGACGGTTGTTGCCCAGCTCAAGGAC 304

Qy         33 GlyGly------------GlySerProThrValHisIleGlyProThrAlaPheLeuGly 48
                 |||            |||         |||   |||   |||
Db        305 AAGGGTTCGGTCAGCCGCGGCTGGATCGGCGTGCAGATTCAGCCGGTGACG--------- 355

Qy         49 LeuGlyValValAspAsnAsnGly------------AsnGlyAlaArgValGlnArgVal 64
                 :::   |||:::   |||                |||||   |||
Db        356 TCGGATATCGCCGACAGCCTCGGCATGAAGAAGGCCGAAGGGGCGCTGGTGGCGGAGCCG 415

Qy         65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
                 :::   |||||||||      ||||||   :::|||||||||||||||:::|||
Db        416 CAGGCGAACGGTCCGGCGGCGAAGGCCGGCATCGAGTCCGGCGACGTGATCACCTCGGTC 475

Qy         85 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 104
              :::|||      :::      |||      :::|||      :::   |||        |||
Db        476 AACGGCGAATCCGTCAAGGACGCCCGCGAGCTCGCCCGCACCATCGGCGGCATGGCGCCC 535

Qy        105 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 124
                 ::::::   :::           |||   |||                  |||:::
Db        536 CGTGCGATCGTGAAGCTTAACGTGCTGCACAAG---GGCCAGGACAATGTCGTGAACCTC 592

Qy        125 ThrLeu 126
              ||||||
Db        593 ACCCTC 598

RESULT 8
BQ514888/c
LOCUS       BQ514888                758 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST622303 Generation of a set of potato cDNA clones for microarray
            analyses mixed potato tissues Solanum tuberosum cDNA clone STMIO81
            3' end, mRNA sequence.
ACCESSION   BQ514888
VERSION     BQ514888.1  GI:21373757
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     Other_ESTs: EST622302
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
```

Seq primer: T7.
FEATURES Location/Qualifiers
source 1. .758
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binjte"
/db_xref="taxon:4113"
/clone="STMIO81"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."

ORIGIN

Alignment Scores:

Pred. No.:	0.449	Length:	758
Score:	113.00	Matches:	34
Percent Similarity:	42.98%	Conservative:	15
Best Local Similarity:	29.82%	Mismatches:	47
Query Match:	17.36%	Indels:	18
DB:	13	Gaps:	3

US-09-684-215B-23 (1-128) x BQ514888 (1-758)

```
Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
Db        592 ACATCAGCAGGTGTTGGATTTGCAATCCCTTCTTCAACTGTGTTGAGAGTTGTGCCCCAG 533

Qy         30 IleArgSerGlyGly------------GlySerProThrValHisIleGlyProThrAla 45
Db        532 TTGATCCAATCTGGAAAAGTTCTTCGTGCTGGTTTGAATATTGAAATCGCTCCAGACCTG 473

Qy         46 PheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValVal 65
Db        472 ---------ATTGCCAACCAACTTAATGTTCGAAATGGAGCACTGGTTTTGCTGGTACCT 422

Qy         66 GlySerAlaProAlaAlaSerLeuGly-------------------------------- 74
Db        421 GGAAATAGTCCTGCAGCGAAAGCGGGACTTCTTCCTACTACCAGGGGTTTTGCAGGAAAT 362

Qy         75 IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAla 94
Db        361 ATAGTGCTTGGCGATATTATTGAAGCAGTGGATGACAAACCTGTTAGGAGTAAAGCAGAG 302

Qy         95 MetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 108
Db        301 TTGTATAAAGCCCTGGATAACTATAACATAGGTGATGAAGTT 260
```

RESULT 9
CD080132
LOCUS CD080132 982 bp mRNA linear EST 14-SEP-2003
DEFINITION MA3-9999U-M294-C03-U.B MA3-0001 Schistosoma mansoni cDNA clone MA3-9999U-M294-C03.B, mRNA sequence.
ACCESSION CD080132
VERSION CD080132.1 GI:34631128
KEYWORDS EST.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE 1 (bases 1 to 982)
AUTHORS Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M., Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr., Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F., Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L., Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A., Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A., Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T., Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M., Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
TITLE Transcriptome analysis of the acoelomate human parasite Schistosoma mansoni
JOURNAL Nat. Genet. 35 (2), 148-157 (2003)
MEDLINE 22879926
COMMENT Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquimica
Instituto de Quimica - Universidade de Sao Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP, Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjo@iq.usp.br
This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL http://bioinfo.iq.usp.br/schisto/
Plate: MA3-9999U-M294 row: 3 column: C.
FEATURES Location/Qualifiers
source 1. .982
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="MA3-9999U-M294-C03.B"
/sex="mixed pool"
/dev_stage="adult"
/lab_host="Mus musculus"
/clone_lib="MA3-0001"

ORIGIN

Alignment Scores:

Pred. No.:	2.05	Length:	982
Score:	108.00	Matches:	33
Percent Similarity:	37.80%	Conservative:	15
Best Local Similarity:	25.98%	Mismatches:	39
Query Match:	16.59%	Indels:	40
DB:	14	Gaps:	4

US-09-684-215B-23 (1-128) x CD080132 (1-982)

```
Qy          6 AsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAla------ 23
Db        355 AACGCCATGGTGGCTGGTACAGGCGTGGGATTTGCTATACCTGTTGATCAAGTCCGAAAA 414

Qy         24 ---------MetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
Db        415 TTCATACAGTTAGCACTAAAAGCATCCTCAAATACTCGGAGCTCTTCCCCAATATCGTCA 474

Qy         41 IleGlyProThr---------------------------------------------- 44
Db        475 AAATCGCCTACTGACCCATATATTTTAGCAGACAATCCAGCCAATTCGGAAAGTGGAACT 534

Qy         45 ------AlaPheLeuGlyLeu--------------------------------GlyVal 51
Db        535 CAACGTCGATACTTAGGCCTAGTTATGCGGACACTTACTCCAGAACTGGCTTTCGAATTA 594

Qy         52 ValAspAsnAsnGly-----------------AsnGlyAlaArgValGlnArgValVal 65
Db        595 GCATCTCGCGGTGGACAACATTTTGTAGAGTTGAATGGTGTTCTGATTCATGCTGTATTA 654

Qy         66 GlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAsp 85
Db        655 AGAAACTCCCCTGCTCANAGAGCGGGATTAAGAGCTGGAGATGTAATTGTTGCCATTGAT 714

Qy         86 GlyAlaProIleAsnSerAla 92
Db        715 GGTCTGCCTATAACGAATGCT 735
```

RESULT 10
BZ561390/c
LOCUS BZ561390 1033 bp DNA linear GSS 17-DEC-2002
DEFINITION pacs2-164_3239.y3 pacs2-164 Pseudomonas aeruginosa genomic clone

```
            pacs2-164_3239, genomic survey sequence.
ACCESSION   BZ561390
VERSION     BZ561390.1  GI:27181349
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .1033
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_3239"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:                3.11          Length:         1033
Score:                    106.50        Matches:        41
Percent Similarity:       44.20%        Conservative:   20
Best Local Similarity:    29.71%        Mismatches:     56
Query Match:              16.36%        Indels:         21
DB:                       28            Gaps:           4

US-09-684-215B-23 (1-128) x BZ561390 (1-1033)

Qy          4 SerAspAsnPheGlnLeuSerGlnGlyGlyGln------GlyPheAlaIleProIleGly 21
              |||     |||     |||   |||             |||   |||||||||
Db        713 TCCCAGTTTTTCCACCGGTCCGCCGGCTCAATGGCCCTGTCTTTCGGCATTCCGATCGAT 654

Qy         22 GlnAlaMetAlaIle-AlaGlyGlnIleArg--------------------------Se 32
                ||:::   ::: |||   |||::::::                          ::
Db        653 GTCGCGCTGACCGTTCGCCGACCAGTTGAAGAAAGCCCGGCAAGGTCAGTCGCGGCTGGC 594

Qy         32 rGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeu-GlyLeuGlyValV 52
              :|||   |||   |||             |||           ||| ||||||
Db        593 TGGGCGTGGTGATCCAGGAAGTGAACAAGGATCTCGCCGAGTCCTTCGGCCTC------- 541

Qy         52 alAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaS 72
                :::     :::||||||   |||   ::::::|||         |||||||||
Db        540 -----GACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCCA 486

Qy         72 erLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerA 92
                |||:::     |||||||||||   :::::::::|||      ||||||   :
Db        485 AGGGTGGCCTGCAGGTGGGCGATGTGATCCTCAGCCTGAACGGCCAGTCGATCAACGAGT 426

Qy         92 laThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal---- 110
              ::      :::          :::           |||||||||   |||::::::
Db        425 CCGCCGACCTGCCGCACCTGGTGGGCAACATGAAGCCGGGCGACAAGATCAACCTGGACG 366

Qy        111 --------ThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 124
                      |||   ||||||                |||::::::
Db        365 TGATTCGCAACGGCCAGCGCAAGTCCCTGAGCATGGCGGTAGGCAGCCTT 316

RESULT 11
```

```
BZ549047/c
LOCUS       BZ549047                947 bp    DNA     linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_1644.s1 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_1644, genomic survey sequence.
ACCESSION   BZ549047
VERSION     BZ549047.1  GI:27152628
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 947)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .947
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="1-60"
                     /db_xref="taxon:287"
                     /clone="pacs1-60_1644"
                     /clone_lib="pacs1-60"
                     /note="clinical isolate 1-60 Whole genomic shotgun
                     library."
ORIGIN

Alignment Scores:
Pred. No.:                5.43          Length:         947
Score:                    103.50        Matches:        37
Percent Similarity:       44.74%        Conservative:   14
Best Local Similarity:    32.46%        Mismatches:     48
Query Match:              15.90%        Indels:         16
DB:                       28            Gaps:           3

US-09-684-215B-23 (1-128) x BZ549047 (1-947)

Qy          4 SerAspAsnPheGlnLeuSerGlnGly------GlyGlnGlyPheAlaIleProIleGly 21
              |||     |||     |||   |||       |||       ||||||||||||||||
Db        374 TCGCAGATCTTCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGAT 315

Qy         22 GlnAlaMetAlaIleAlaGlyGlnIleArgSer--------------------------- 32
                ||||||   :::|||   |||::::::
Db        314 GTCGCGATGAACGTCGCCGACCAGTTGAAGAAAGCCGGCAAGGTCAGTCGCGGCTGGCTG 255

Qy         33 GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal 52
              |||   |||   |||                     ||||||
Db        254 GGCGT-GGTGATCCAGGAAGTGAACAACGATCTCGCCGAGTCTTCGGCCTC--------- 205

Qy         53 AspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSer 72
                :::     :::||||||   |||   ::::::|||         |||||||||
Db        204 ---GACAAGCCGTCCGGCGCGCTGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCGAAG 148

Qy         73 LeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
              |||:::     |||||||||||   :::::::::|||      ||||||   :::
Db        147 GGTGGCCTGCAGGTGGGCGATGTGATCAACAGCCTGAACGGCCAGTCGATCAACGAGTCC 88

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAsp 106
                ::      :::          :::           |||||||||
Db         87 GCCGACCTGCCGCACCTGGTGTGGAACATGAAGCCGGGCGAC 46

RESULT 12
AQ989479
```

```
LOCUS       AQ989479                726 bp    DNA     linear   GSS 14-AUG-2000
DEFINITION  Rfc00025 Photorhabdus luminescens strain W14 M13 library
            Photorhabdus luminescens genomic clone PLG00025, genomic survey
            sequence.
ACCESSION   AQ989479
VERSION     AQ989479.1  GI:9648073
KEYWORDS    GSS.
SOURCE      Photorhabdus luminescens
  ORGANISM  Photorhabdus luminescens
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Photorhabdus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   ffrench-Constant,R.H., Waterfield,N., Burland,V., Perna,N.T.,
            Daborn,P.J., Bowen,D. and Blattner,F.R.
  TITLE     A genomic sample sequence of the entomopathogenic bacterium
            Photorhabdus luminescens W14: potential implications for virulence
  JOURNAL   Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)
  MEDLINE   20378633
   PUBMED   10919786
COMMENT     Contact: ffrench-Constant RH
            Department of Biology and Biochemistry
            University of Bath
            South Building, Bath BA2 7AY, UK
            Tel: (44) 1225 826621
            Fax: (44) 1225 826779
            Email: bssrfc@bath.ac.uk
            This is one of 2,122 random reads from the M13 library. For
            annotation of identified clones (BLASTX, BLASTN and mapping to E.
            coli K12 genome) please see ffrench-Constant et al. 2000, Nucleic
            Acids Res.
            Seq primer: M13 Forward
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Photorhabdus luminescens"
                     /mol_type="genomic DNA"
                     /strain="W14"
                     /db_xref="taxon:29488"
                     /clone="PLG00025"
                     /dev_stage="primary phase variant"
                     /clone_lib="Photorhabdus luminescens strain W14 M13
                     library"
                     /note="Genomic DNA from strain W14 was size selected (1-2
                     kb) and then cloned into M13 Janus."
ORIGIN
```

Alignment Scores:

Pred. No.:	4.14	Length:	726
Score:	103.00	Matches:	39
Percent Similarity:	46.55%	Conservative:	15
Best Local Similarity:	33.62%	Mismatches:	52
Query Match:	15.82%	Indels:	11
DB:	28	Gaps:	2

US-09-684-215B-23 (1-128) x AQ989479 (1-726)

```
Qy          6 AsnPheGlnLeuSerGlnGlyGlyGln----------GlyPheAlaIleProIle 20
              :::|||     |||:::  ||:::               ||||||||||||||
Db        179 TCATTCGATAAAAGTGAAAATGGTGAAACGCCAGAAGGGTTGGGGTTCGCTATTCCGACT 238

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                  |||      |||    ::::::     |||
Db        239 GAACTCGCTACTAAGATTATGCAGAAACTTATCCGTGACGGACGAGTTATCCGGGGTTTT 298

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn--------- 57
              ||||||   |||:::          :::     :::||||||||||
Db        299 ATCGGCATAACATCGCAAGAGCTGCCACATATTCGCTCTTCTAACGGCAATATTAACCAG 358

Qy         58 ------GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIle 75
                   |||   ||||||   :::|||      :::   ||||||          ||||||
Db        359 CTTCAAGGGCTACGAGTATTTCAGGTAACTACAAACGGGCCTGCGCAAAAAGCCGGTATT 418

Qy         76 SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet 95
```

```
                  ||||||:::||||||:::|||:::      |||      ||||||
Db        419 AAAGTGGGTGATATCATTACCAGCGTCAATAACAAACCCGCTATTTCTGCGGGTGAAACA 478

Qy         96 AlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 111
                 |||   :::      :::   ||||||  |||      ||||||
Db        479 ATGGATCAAGTGGCTGA-AATACCCCCGGCAGCGTGGGTNCTGTCACT 525
```

```
RESULT 13
AZ933900
LOCUS       AZ933900                719 bp    DNA     linear   GSS 24-APR-2001
DEFINITION  BJ__Ba0001B03r B. japonicum BAC library Bradyrhizobium japonicum
            genomic, genomic survey sequence.
ACCESSION   AZ933900
VERSION     AZ933900.1  GI:13775960
KEYWORDS    GSS.
SOURCE      Bradyrhizobium japonicum
  ORGANISM  Bradyrhizobium japonicum
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1  (bases 1 to 719)
  AUTHORS   Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A.,
            Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
  TITLE     A marker-dense, sequence-ready map of the Bradyrhizobium japonicum
            genome
  JOURNAL   Genome Res. 11 (8), 1434-1440 (2001)
  MEDLINE   21376150
   PUBMED   11483585
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Class: BAC ends
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .719
                     /organism="Bradyrhizobium japonicum"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:375"
                     /lab_host="E. coli"
                     /clone_lib="B. japonicum BAC library"
                     /note="Vector: pIndigo536; Site_1: HindIII"
ORIGIN
```

Alignment Scores:

Pred. No.:	5.13	Length:	719
Score:	102.00	Matches:	26
Percent Similarity:	44.44%	Conservative:	14
Best Local Similarity:	28.89%	Mismatches:	40
Query Match:	15.67%	Indels:	10
DB:	28	Gaps:	2

US-09-684-215B-23 (1-128) x AZ933900 (1-719)

```
Qy         13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
              |||   ||||||||||||   :::                :::          :::
Db         44 GGCATCGGCTTTGCGATCCTTGTCAACATGGAGCGCGTGGTCGTCGCTTCCGCCAAGGGC 103

Qy         33 GlyGly-----------------------GlySerProThrValHisIleGlyProThr 44
              ||||||                      |||:::           :::   |||
Db        104 GGCGGCAAGGCGGTGAAGCGTCCCTGGCTCGGCGCGAAGTTGCAGGCGGTGACGCCCGAG 163

Qy         45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64
                     ||||||:::      :::     ||||||  |||     |||
Db        164 ATCGCCGAAAGCCTCGGCTTG------CGTTCGCCGACCGGCGCGCTGGTCGCGAGCGTG 217

Qy         65 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
              |||   :::   |||||||||      |||:::   :::   |||:::||||||   :::
Db        218 GTCTCGAACGGTCCCGCAGCGAAGGCCGGCCTGAAATCCTCCGATCTGATCACCGGGATC 277
```

```
Qy         85 AspGlyAlaProIleAsnSerAlaThrAla 94
              ||||||      ::::::      |||
Db        278 GACGGCCAGACCGTGGATGATCCCAATGCC 307

RESULT 14
AJ558965
LOCUS       AJ558965                765 bp    mRNA    linear   EST 12-JUN-2003
DEFINITION  AJ558965 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_1_09_a08, mRNA sequence.
ACCESSION   AJ558965
VERSION     AJ558965.1  GI:31661537
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Lamiales; Antirrhinaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.
  TITLE     Antirrhinum EST collection
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Antirrhinum majus"
                     /mol_type="mRNA"
                     /db_xref="taxon:4151"
                     /clone="018_1_09_a08"
                     /tissue_type="whole plant"
                     /clone_lib="Antirrhinum majus whole plant"
ORIGIN

Alignment Scores:
Pred. No.:                5.61          Length:          765
Score:                    102.00        Matches:         34
Percent Similarity:       40.83%        Conservative:    15
Best Local Similarity:    28.33%        Mismatches:      41
Query Match:              15.67%        Indels:          30
DB:                       9             Gaps:            4

US-09-684-215B-23 (1-128) x AJ558965 (1-765)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
               :::   |||    ||||||||||||||||          :::   |||      |||
Db         311 ACCTCGGCGGGTGTAGGATTTGCGATCCCTTCATCGACAGTGGTAAAGATAGTGCCTCAG 370

Qy          30 IleArgSerGlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeu 49
               :::                  ::::::   |||                  ||||||
Db         371 TTG----------------------ATCAATTATGGAAAAGTTGTTCGAGCTGGTTTA 406

Qy          50 GlyValValAsp----------------------------AsnAsnGlyAsnGlyAla 59
                  ||||||                            |||     |||||||||
Db         407 AAT---GTTGACATAGCTCCTGATCTTATTGCAAATCAACTCAATGTTCGAAATGGAGCT 463

Qy          60 ArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGly-------------- 74
                  :::   :::|||   |||::::::   ||||||      |||
Db         464 CTAATCTTGCAGGTCCCTGGAAATAGTTCAGCAGCTAAGGCTGGGCTTGTACCTACAACA 523

Qy          75 ------------------IleSerThrGlyAspValIleThrAlaValAspGlyAlaPro 88
                               |||      ||||||:::|||   |||||||||      |||
Db         524 CGGGGTTTCGCTGGTAAAATCGTACTTGGGGATATAATTTTGGCCGTAGATGACAAACCT 583

Qy          89 IleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIle 108
               :::   |||         :::         |||:::   :::      ||||||   :::
Db         584 GTCAGGAGTAGGGCCGAGCTATACAAAACACTGGATGAGTATTCCATCGGAGACAAAGTG 643

RESULT 15
CA480614/c
LOCUS       CA480614                859 bp    mRNA    linear   EST 13-NOV-2002
DEFINITION  AGENCOURT_10739220 NCI_CGAP_ZEmb2 Danio rerio cDNA clone
            IMAGE:6789351 5', mRNA sequence.
ACCESSION   CA480614
VERSION     CA480614.1  GI:24938284
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 859)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Leonard I. Zon, M.D.
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14289  row: k  column: 14
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .859
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6789351"
                     /tissue_type="embryo"
                     /lab_host="DH10B (T1-resistant)"
                     /clone_lib="NCI_CGAP_ZEmb2"
                     /note="Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 2 kb. Constructed by J. Wang (Research Genetics,
                     Invitrogen Corp) from tissue donated by L. Zon (Harvard
                     University). Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:                9.36          Length:          859
Score:                    100.50        Matches:         36
Percent Similarity:       40.30%        Conservative:    18
Best Local Similarity:    26.87%        Mismatches:      55
Query Match:              15.44%        Indels:          25
DB:                       14            Gaps:            5

US-09-684-215B-23 (1-128) x CA480614 (1-859)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle--------- 26
               :::   |||          |||:::::::|||   |||       :::::::::
Db         854 GCCTCTGGGCACCTGAACTTCTCCTTGCCCAATGGTGATGACTTGTCCGTCAGGCAGCTC 795

Qy          27 AlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle--------------- 41
                      ::::::   ||||||||||||          |||:::
Db         794 GTAGCTCTTCTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCATCTCCTGCTCAAAGTCAAG 735

Qy          42 GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgVal 61
               |||       |||   |||   |||   |||   ||||||
Db         734 GGCCACATAGCAGAGCTTCTCCTTGATGTCACGGACAAT-----------------TTC 693

Qy          62 GlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIle 81
                  |||              ||||||||||||              :::|||:::
Db         692 CCTCTCAGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAGGATCTTCATGAGGTAGTC 633

Qy          82 ThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla--------- 98
               :::   |||      :::   :::   :::      |||
Db         632 AGTCAGGTCACGGCCAGCCAAGTCCAGACGGAGGATGGCATGGGGCAGGGCGTAACCCTC 573

Qy          99 LeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGly 118
               ::::::||||||      ||||||   |||:::                  ||||||
```

```
Db        572 GTAGATGGGCACAGTGTGGGTGACACCATCACCAGAGTCCATCACGATACCAGTGGTACG 513

Qy        119 ThrArg----------------------ThrGlyAsnVal 124
              ||||||                      ||||||||
Db        512 ACCAGAGGCATACAGGGACAGCACAGCCTGGATGGCAACGTA 471
```

RESULT 16
AI597611

LOCUS	AI597611 423 bp mRNA linear EST 21-APR-1999
DEFINITION	tn15f02.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2167707 3' similar to TR:O08664 O08664 BCL7C MRNA. ;, mRNA sequence.
ACCESSION	AI597611
VERSION	AI597611.1 GI:4606659
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 423)
AUTHORS	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL	Unpublished (1998)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Seq primer: -40UP from Gibco.

FEATURES Location/Qualifiers

source 1. .423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2167707"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn25"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.:	4.71	Length:	423
Score:	99.00	Matches:	35
Percent Similarity:	44.35%	Conservative:	20
Best Local Similarity:	28.23%	Mismatches:	43
Query Match:	15.21%	Indels:	26
DB:	9	Gaps:	7

US-09-684-215B-23 (1-128) x AI597611 (1-423)

```
Qy          7 PheGlnLeuSerGln---------GlyGlyGlnGlyPheAlaIleProIleGlyGlnAla 23
              |||:::  ||||||          |||                 |||:::  |||
Db         39 TTTGAAAAAAGCCAAAGGGGCCCCTGGGGCAACAGGACAGGCAGGCCGGTTTCTCAGGGG 98

Qy         24 MetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThr---ValHisIleGly 42
                        ||||||:::|||      ||||||:::|||:::        :::
Db         99 TCAGGGGCATTTGGGCAGATGCGCTTGAGTGGGGGGGCACCCTCCAAGTCCTTTGTGTCA 158

Qy         43 ProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGln 62
              |||  |||     |||  |||                 |||  |||::::::
Db        159 CCCTGGGCTGCCTCAGGGACAGGT-------------GGCACTGGCTCAAAAACAGGG 203

Qy         63 ArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThr 82
                  |||::::::|||::::::|||  |||      ||||||      :::
Db        204 TAAGCTTCGGGGGCCTCACCTTCCAGCAGTTCTGGA------ACAGGCTCCTCCTTGGTC 257

Qy         83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
              ::::::  |||:::  :::     |||
Db        258 AGCATTGGGGGTTCGTCGGTGCTGCCAGCAGTTATG---------------------- 293

Qy        103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGly-----------Gly 118
                |||||         |||:::  |||  |||:::  |||           |||
Db        294 CCCCCGGGA---------TCTCTTTTTGGCCCACCCGTGGGGGCTGAGCCTTCTCAGGG 344

Qy        119 ThrArgThrGly 122
              |||  :::|||
Db        345 ACCCCTTCTGGG 356
```

RESULT 17
BI350520/c

LOCUS	BI350520 499 bp mRNA linear EST 26-JUL-2002
DEFINITION	fr32b05.y1 zebrafish adult brain Danio rerio cDNA clone IMAGE:4955073 5' similar to SW:ACT2_FUGRU P53485 ACTIN, CYTOPLASMIC 2 ;, mRNA sequence.
ACCESSION	BI350520
VERSION	BI350520.1 GI:15044966
KEYWORDS	EST.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE	1 (bases 1 to 499)
AUTHORS	Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE	WashU Zebrafish EST Project 1998
JOURNAL	Unpublished (1998)
COMMENT	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbrafish@watson.wustl.edu cDNA Library Preparation: John Ngai. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de) Seq primer: T7 High quality sequence stop: 301.

FEATURES Location/Qualifiers

source 1. .499
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:4955073"
/sex="mixed male and female"
/tissue_type="brain"
/dev_stage="adult"
/lab_host="E. coli DH10B"
/clone_lib="zebrafish adult brain"
/note="Vector: pZIPLOX; Site_1: NotI; Site_2: SalI; Original library was constructed in lambdaZIPLOX. Mass

excision of the cDNA library was performed to yield pZIPLOX plasmids. Insert check was done in original library."

ORIGIN

Alignment Scores:

Pred. No.:	8.44	Length:	499
Score:	97.50	Matches:	38
Percent Similarity:	39.29%	Conservative:	17
Best Local Similarity:	27.14%	Mismatches:	52
Query Match:	14.98%	Indels:	33
DB:	12	Gaps:	6

US-09-684-215B-23 (1-128) x BI350520 (1-499)

```
Qy          9 LeuSerGlnGly----------------------GlyGlnGlyPheAlaIleProIle 20
              |||   ||||||                      |||      :::   :::
Db        430 CTGGAACAGGGCCTCTGGACACCTGAACCTTTCATTGCCAATGGTGATGACCTGTCCGTC 371

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                 |||   :::|||:::        ::::::   ||||||||||||         |||
Db        370 AGGCAGCTGGTAGCTCTT---------CTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCAT 320

Qy         41 Ile---------------GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsn 55
              :::              |||      |||   |||   |||   |||   ||||||
Db        319 CTCCTGCTCAAAGTCAAGGGCCACATAGCAGAGCTTCTCCTTGATGTCACGGACAAT--- 263

Qy         56 GlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIle 75
                                          |||         ||||||||||||
Db        262 ---------------TTCCCTCTCAGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAG 218

Qy         76 SerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMet 95
                    :::|||::::::   |||      :::   :::   :::          |||
Db        217 GATCTTCATGAGGTAGTCAGTCAGGTCACGGCCAGCCAAGTCCAGACGGAGGATGGCATG 158

Qy         96 AlaAspAla---------LeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112
                             ::::::|||||||      ||||||   |||::::::
Db        157 GGGCAGGGCGTAACCCTCGTAGATGGGCACAGTGTGGGTGACACCATCACCATAGTCCAT 98

Qy        113 GlnThrLysSerGlyGlyThrArg----------------------ThrGlyAsnVal 124
                  ||||||   ||||||                           |||||||||
Db         97 CACGATACCAGTGGTACGACCAGAGGCATACAGGAACAGCACAGCCTGGATGGCAACGTG 38
```

RESULT 18
CA473840/c

```
LOCUS       CA473840               870 bp    mRNA    linear   EST 12-NOV-2002
DEFINITION  AGENCOURT_10740457 NCI_CGAP_ZKid1 Danio rerio cDNA clone
            IMAGE:6794807 5', mRNA sequence.
ACCESSION   CA473840
VERSION     CA473840.1  GI:24930192
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 870)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Leonard I. Zon, M.D.
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14303  row: n  column: 22
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .870
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6794807"
                     /lab_host="DH10B (T1-resistant)"
                     /clone_lib="NCI_CGAP_ZKid1"
                     /note="Organ: kidney; Vector: pCMV-SPORT6.1; Site_1:
                     EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 1.8 kb. Constructed by J.
                     Wang (Research Genetics, Invitrogen Corp) from tissue
                     donated by L. Zon (Harvard University). Note: this is a
                     NCI_CGAP Library."
ORIGIN
```

Alignment Scores:

Pred. No.:	18.9	Length:	870
Score:	97.50	Matches:	35
Percent Similarity:	40.62%	Conservative:	17
Best Local Similarity:	27.34%	Mismatches:	51
Query Match:	14.98%	Indels:	25
DB:	14	Gaps:	5

US-09-684-215B-23 (1-128) x CA473840 (1-870)

```
Qy         16 PheAlaIleProIleGlyGlnAlaMetAlaIle---------AlaGlyGlnIleArgSer 32
              |||::::::|||   |||      :::::::::                 ::::::
Db        811 TTTTCATTGCCCAATGGTGATGACCTGTCCGTCAGGCAGCTCGTAGCTCTTCTCCAGGGA 752

Qy         33 GlyGlyGlySerProThrValHisIle--------------GlyProThrAlaPheLeu 47
              ||||||||||||      |||:::              |||    |||   |||
Db        751 GGAGGAGGAAGCAGCGGTGCCCATCTCCTGCTCAAAGTCAAGGGCCACATAGCAGAGCTT 692

Qy         48 GlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySer 67
              |||   |||   ||||||                                  |||
Db        691 CTCCTTGATGTCACGGACAAT-----------------TTCCCTCTCAGCTGTGGTGGT 650

Qy         68 AlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAla 87
                    ||||||||||||           :::|||::::::   |||      :::
Db        649 GAAGCTGTAGCCTCTCTCGGTCAGGATCTTCATGAGGTAGTCAGTCAGGTCACGGCCAGC 590

Qy         88 ProIleAsnSerAlaThrAlaMetAlaAspAla---------LeuAsnGlyHisHisPro 104
                 :::   :::          |||                 ::::::||||||
Db        589 CAAGTCCAGACGGAGGATGGCATGGGGCAGGGCGTAACCCTCGTAGATGGGCACAGTGTG 530

Qy        105 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg------------ 120
              ||||||   |||:::              |||||   ||||||
Db        529 GGTGACACCATCACCAGAGTCCATCACGATACCAGTGGTACGACCAGAGGCATACAGGGA 470

Qy        121 ------------ThrGlyAsnVal 124
                         |||||||||
Db        469 CAGCACAGCCTGGATGGCAACGTA 446
```

RESULT 19
CA148171

```
LOCUS       CA148171               444 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  SCEZRZ1017F10.g RZ1 Saccharum officinarum cDNA clone SCEZRZ1017F10
            5', mRNA sequence.
ACCESSION   CA148171
VERSION     CA148171.1  GI:35048941
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
```

```
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 017  row: F  column: 10
            Seq primer: T7 Promoter Primer.
FEATURES              Location/Qualifiers
     source           1. .444
                      /organism="Saccharum officinarum"
                      /mol_type="mRNA"
                      /db_xref="taxon:4547"
                      /clone="SCEZRZ1017F10"
                      /lab_host="DH10B"
                      /clone_lib="RZ1"
                      /note="Organ: Shoot-root transition zone from young plants
                      (large insert library); Vector: pSport1; Site_1: SalI;
                      Site_2: NotI; An unidirectional cDNA library generated
                      from [Shoot-root transition zone  from young plants (large
                      insert library)]. cDNA was prepared from polyA+  mRNA
                      using SuperScript Plasmid System Kit (Invitrogen). The
                      double-strand  cDNAs were fractionated in a sepharose
                      CL-2B 40cm-columns and fragments  sizing between 0.8 and
                      1.5 Kb were directionally cloned into the vector.  Details
                      of each source of RNA and library construction can be
                      obtained at  http://sucest.lad.ic.unicamp.br/public"
ORIGIN

Alignment Scores:
Pred. No.:                7.99          Length:         444
Score:                    97.00         Matches:        26
Percent Similarity:       47.62%        Conservative:   14
Best Local Similarity:    30.95%        Mismatches:     26
Query Match:              14.90%        Indels:         18
DB:                       13            Gaps:           2

US-09-684-215B-23 (1-128) x CA148171 (1-444)

Qy          54 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 73
               |||       |||||||||   :::   :::|||   |||   :::   ||||||
Db          32 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAGGCA 91

Qy          74 GlyIleSer--------------------------------ThrGlyAspValIleThr 82
               |||::::::                               ||||||||||||
Db          92 GGCCTTGCTCCGACCGGCAGGGGTTTCGCTGGTAACATTGTTCTTGGTGATGTCATCGTT 151

Qy          83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
               ||||||||||||  |||:::       :::   :::       |||:::   :::
Db         152 GCAGTGGATGGCAAACCTGTTAAGGGCAAATCAGACCTGCTGAGGGTTCTGGATGACTAT 211

Qy         103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122
                  ||||||   :::::::::|||                          |||   |||
Db         212 GGCGTCGGGGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 250

Qy         123 AsnValThrLeu 126
               :::  ||||||
Db         251 TCTGAAACCCTC 262

RESULT 20
BI721127
LOCUS       BI721127                543 bp    mRNA    linear   EST 19-SEP-2001
DEFINITION  1031054B10.y1 C. reinhardtii CC-1690, Stress II (normalized),
            Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BI721127
VERSION     BI721127.1  GI:15696822
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 543)
  AUTHORS   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
            Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1031
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES              Location/Qualifiers
     source           1. .543
                      /organism="Chlamydomonas reinhardtii"
                      /mol_type="mRNA"
                      /strain="CC-1690 wild type mt+ 21gr"
                      /db_xref="taxon:3055"
                      /clone_lib="C. reinhardtii CC-1690, Stress II
                      (normalized), Lambda Zap II"
                      /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                      XhoI; Stress condition II library, constructed by John
                      Davies and Jeffrey McDermott, combines cDNAs from CC-1690
                      cells grown to mid-log phase in TAP (NH4+ - containing)
                      and shifted to TAP - NO3- (24hrs); H2 production
                      conditions (0, 12hr, 24hr) see Melis et al.,(2000) Plant
                      Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
                      sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
                      PolyA mRNA was purified from each sample, pooled and cDNA
                      synthesized. The cDNA was directionally cloned into lambda
                      Zap II (Stratagene) in the EcoRI (5') and XhoRI (3')
                      sites. pBluescript II SK- plasmids were excised from the
                      lambda ZAP clones by superinfection with ExAssist
                      (Stratagene) phage. The library was normalized using
                      method 4 described in Bonaldo et al., (1996) Genome
                      Research 6: 791-806."
ORIGIN

Alignment Scores:
Pred. No.:                10.7          Length:         543
Score:                    97.00         Matches:        31
Percent Similarity:       39.83%        Conservative:   16
Best Local Similarity:    26.27%        Mismatches:     49
Query Match:              14.90%        Indels:         22
DB:                       12            Gaps:           4

US-09-684-215B-23 (1-128) x BI721127 (1-543)

Qy          10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGln 29
               :::     |||  |||||||||:::||||||                 :::     |||
Db         153 GCATCCTCCGGCGTGGGCTTCGCGCTGCCCATCGACAGTGTGCGCGGCCTGGTAGACCAG 212

Qy          30 IleArgSerGlyGlyGly------SerProThrValHisIleGly---------ProThr 44
               |||  :::  |||           |||   :::   :::           |||
Db         213 ATCCTGACCTACGGCCGCGTGCTGCGGCCCGTGCTGGGAGTCAGCCTGGCGCCGCCGCAG 272

Qy          45 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 64
                    ||||||           |||     |||     |||         |||
Db         273 GTGCTCAAGCAGCTGGGC--------------CAGCCGGGCGTGCTGGTGCTAGAGGTG 317

Qy          65 ValGlySerAlaProAlaAlaSerLeuGlyIleSer---------------------- 76
                  :::||||||         ||||||
Db         318 CCCAAGGGCTCCCCCGCGGAGAAGGCCGGCATCAAGCCCACCATGCGCGACCGCTTCAGC 377

Qy          77 -----------ThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
                         ||||||:::||||||   :::||||||      :::   :::
Db         378 GGCTCCCTGGTGCTGGGCGACATCATCACAGGCATTGACGGCAAGGCCGTAAAGAACTAT 437

Qy          93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
               :::   :::   :::||||||:::          ||||||   :::   |||
Db         438 TCGGACCTGGTGGAGGCGCTGGATGAGAAGCGCGTGGGCGACACAGTCAAGGTG 491

RESULT 21
```

```
AW285510
LOCUS       AW285510                546 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  LG1_241_E05.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW285510
VERSION     AW285510.1  GI:6675354
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 546)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence start: 126
            High quality sequence stop: 546
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .546
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

Alignment Scores:
Pred. No.:                 13.5          Length:          546
Score:                     96.00         Matches:         25
Percent Similarity:        47.62%        Conservative:    15
Best Local Similarity:     29.76%        Mismatches:      26
Query Match:               14.75%        Indels:          18
DB:                        10            Gaps:            2

US-09-684-215B-23 (1-128) x AW285510 (1-546)

Qy          54 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 73
               |||       |||||||||  :::   :::|||   |||   :::   ||||||
Db          56 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAAGCA 115

Qy          74 GlyIleSer--------------------------------ThrGlyAspValIleThr 82
               |||::::::                               ||||||:::|||
Db         116 GGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTGGGTGATATCATCGTT 175

Qy          83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
               ||||||||||||   |||:::        :::   :::         |||:::   :::
Db         176 GCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGGGTTCTGGATGACTAT 235

Qy         103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122
               ||||||   :::::::::|||                                |||   |||
Db         236 GGCGTCGGAGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 274

Qy         123 AsnValThrLeu 126
               :::   ||||||
Db         275 TCAGAAACCCTT 286
```

```
RESULT 22
AW285527
LOCUS       AW285527                551 bp    mRNA    linear   EST 19-JUL-2000
DEFINITION  LG1_241_G05.g1_A002 Light Grown 1 (LG1) Sorghum bicolor cDNA, mRNA
            sequence.
ACCESSION   AW285527
VERSION     AW285527.1  GI:6675371
KEYWORDS    EST.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 551)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.
  TITLE     An EST database from Sorghum: light-grown seedlings
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Laboratory for Genomics and Bioinformatics
            The University of Georgia, Department of Plant Biology
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 583 0210
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: T7
            High quality sequence start: 16
            High quality sequence stop: 552
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .551
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /db_xref="taxon:4558"
                     /clone_lib="Light Grown 1 (LG1)"
                     /note="Organ: 10- to 14-day-old light-grown (greenhouse)
                     seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2:
                     EcoRI; The library was made from poly-A RNA in the cloning
                     vector lambda ZAP II.  Clones to be sequenced were
                     prepared by mass excision."
ORIGIN

Alignment Scores:
Pred. No.:                 13.7          Length:          551
Score:                     96.00         Matches:         25
Percent Similarity:        47.62%        Conservative:    15
Best Local Similarity:     29.76%        Mismatches:      26
Query Match:               14.75%        Indels:          18
DB:                        10            Gaps:            2

US-09-684-215B-23 (1-128) x AW285527 (1-551)

Qy          54 AsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeu 73
               |||       |||||||||  :::   :::|||   |||   :::   ||||||
Db          61 AATGTTCGCAACGGAGCTCTTATACTTAAGGTACCTGGGGGCAGTGCTGCAGCCAAAGCA 120

Qy          74 GlyIleSer--------------------------------ThrGlyAspValIleThr 82
               |||::::::                               ||||||:::|||
Db         121 GGTCTTGCTCCAACCGGCAGGGGTTTTGCTGGTAATATTGTTCTGGGTGATATCATCGTT 180

Qy          83 AlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHis 102
               ||||||||||||   |||:::        :::   :::         |||:::   :::
Db         181 GCAGTGGACGGCAAACCTGTTAAGGGCAAATCTGACCTGCTGAGGGTTCTGGATGACTAT 240

Qy         103 HisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGly 122
               ||||||   :::::::::|||                                |||   |||
Db         241 GGCGTCGGAGATCAGGTGACCTTGACA--------------------ATCCGGCGAGGC 279

Qy         123 AsnValThrLeu 126
               :::   ||||||
Db         280 TCAGAAACCCTT 291
```

```
RESULT 23
BI378928
LOCUS       BI378928                677 bp    mRNA    linear   EST 26-AUG-2003
DEFINITION  BFLG1_000559 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG
            or MPMGp498) Branchiostoma floridae cDNA clone MPMGp498P1235 5',
            mRNA sequence.
ACCESSION   BI378928
VERSION     BI378928.1  GI:30914099
KEYWORDS    EST.
SOURCE      Branchiostoma floridae (Florida lancelet)
  ORGANISM  Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J.,
            Herwig,R., Vingron,M. and Lehrach,H.
  TITLE     New evidence for genome-wide duplications at the origin of
            vertebrates using an amphioxus gene set and completed animal
            genomes
  JOURNAL   Genome Res. 13 (6A), 1056-1066 (2003)
  MEDLINE   22683279
   PUBMED   12799346
COMMENT     Contact: Panopoulou G
            laboratory 145, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1235
            Fax: +49 30 8413 1128
            Email: panopoul@molgen.mpg.de
            The library was characterised by oligonucleotide fingerprinting
            (ONFP) to reduce sequencing redundancy. According to the ONFP
            procedure, clones giving the same hybridisation pattern with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per cluster is selected for sequencing. The size of each
            cluster is an indicator of the frequency of a transcript in the
            analysed library. The cluster size as well the coordinates of the
            rest of the clones assigned to the same fingerprint cluster as the
            clone from which the above EST is generated is available at the
            amphioxus project site at: http://www.molgen.mpg.de/amphioxus/
            Clones and filters are distributed via the Resource Center/Primary
            Database of the German Genome Project (http://www.rzpd.de)
            PCR PRimers
            FORWARD: 5' CCCCAGGCTTTACACTTTATGCTTCCGGCTCG 3' (M13RSP)
            BACKWARD: 5' GCTATTACGCCAGCTGGCGAAAGGGGGATGTG 3' (M13FSP)
            Insert Length: 1200   Std Error: 0.00
            Seq primer: 5'-CCGGTCCGGAATTCCCGGGT-3' pSport3/86
            High quality sequence stop: 677.
FEATURES              Location/Qualifiers
     source           1. .677
                      /organism="Branchiostoma floridae"
                      /mol_type="mRNA"
                      /db_xref="taxon:7739"
                      /clone="MPMGp498P1235"
                      /tissue_type="whole embryo"
                      /dev_stage="5-6 hrs (gastrula stage)"
                      /lab_host="E.coli, XL1 blue"
                      /clone_lib="Amphioxus 5-6 hrs cDNA library (Name
                      convention: BFLG or MPMGp498)"
                      /note="Vector: pSport1; Site_1: SalI, KpnI, EcoRI (5');
                      Site_2: NotI, BamHI, HindIII (3'); OligodT primed and
                      directionally cloned in pSport1 vector using a  NotI
                      (5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC (T)15-3' and a SalI 5'-
                      TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."
ORIGIN

Alignment Scores:
Pred. No.:                20.7          Length:         677
Score:                    95.50         Matches:        21
Percent Similarity:       61.11%        Conservative:   12
Best Local Similarity:    38.89%        Mismatches:     18
Query Match:              14.67%        Indels:         3
DB:                       12            Gaps:           1

US-09-684-215B-23 (1-128) x BI378928 (1-677)

Qy          58 GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThr 77
               |||       |||   |||::::::      :::||||||   |||   ||||||
Db         143 GGAGTACTGGTGCACAGGATCATTGTGGGGTCACCTGCATACAGTGCGGGTATCCGTCCT 202

Qy          78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
               ||||||||||||||:::::::::|||       :::   ||||||      :::   |||
Db         203 GGTGACGTCATCACCAGTATCAACGGTCGATAGGTGACGTCAGCACGAGACATCTACGAC 262

Qy          98 AlaLeuAsnGlyHisHisProGlyAspValIleSerValThr 111
               |||:::|||           |||      :::::::::|||
Db         263 GCGGTAAATAGC---------GGGCAGCAGCTGAACATCACC 295

RESULT 24
CA975497/c
LOCUS       CA975497                909 bp    mRNA    linear   EST 06-JAN-2003
DEFINITION  AGENCOURT_11067943 NCI_CGAP_ZKid1 Danio rerio cDNA clone
            IMAGE:6805289 5', mRNA sequence.
ACCESSION   CA975497
VERSION     CA975497.1  GI:27508151
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 909)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Leonard I. Zon, M.D.
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14331  row: c  column: 16
            High quality sequence start: 10
            High quality sequence stop: 695.
FEATURES              Location/Qualifiers
     source           1. .909
                      /organism="Danio rerio"
                      /mol_type="mRNA"
                      /db_xref="taxon:7955"
                      /clone="IMAGE:6805289"
                      /lab_host="DH10B (T1-resistant)"
                      /clone_lib="NCI_CGAP_ZKid1"
                      /note="Organ: kidney; Vector: pCMV-SPORT6.1; Site_1:
                      EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
                      Oligo dT. Average insert size 1.8 kb. Constructed by J.
                      Wang (Research Genetics, Invitrogen Corp) from tissue
                      donated by L. Zon (Harvard University). Note: this is a
                      NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:                35.7          Length:         909
Score:                    95.00         Matches:        38
Percent Similarity:       39.86%        Conservative:   17
Best Local Similarity:    27.54%        Mismatches:     55
Query Match:              14.59%        Indels:         28
DB:                       14            Gaps:           6

US-09-684-215B-23 (1-128) x CA975497 (1-909)

Qy           9 LeuSerGlnGlyGlyGlnGly---------PheAlaIleProIleGlyGlnAlaMetAla 25
               |||   ||||||      |||          |||   :::|||   |||      ::::::
Db         808 CTGGAACAGGGGCCTCTGGGGCACTTGAACTTCTTCATGCCCAATGGTGATGACCTGTCC 749
```

```
Qy          26 Ile---------AlaGlyGlnIleArgSerGlyGlyGlySerProThrValHisIle--- 41
               :::             ::::::   ||||||||||||        |||:::
Db         748 GTCAGGCAGCTCGTAGCTCTTCTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCATCTCCTG 689

Qy          42 ------------GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsn 57
                          |||     |||   |||   |||   |||   ||||||
Db         688 CTCAAAGTCAAGGGCCACATAGCAGAGCTTCTCCTTGATGTCACGGACAAT--------- 638

Qy          58 GlyAlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThr 77
                                    |||           ||||||||||
Db         637 ---------TTCCCTCTCAGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAGGATCTT 587

Qy          78 GlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAsp 97
               :::|||::::::   |||   :::   :::   :::         :::
Db         586 CATGAGGTAGTCAGTCAGGTCACGGCCAGCCAAGTCCAGACGGAGGATGGCGTGGGGCAG 527

Qy          98 Ala---------LeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThr 114
                        ::::::||||||     ||||||   |||:::
Db         526 GGCGTAACCCTCGTAGATGGGCACAGTGTGGGTGACACCATCACCAGAGTCCATCACGAT 467

Qy         115 LysSerGlyGlyThrArg----------------------ThrGlyAsnVal 124
               ||||||   ||||||                        |||||||||
Db         466 ACCAGTGGTACGACCAGAGGCATACAGGGACAGCACAGCCTGGATGGCAACGTA 413
```

RESULT 25
CF787130/c

LOCUS	CF787130 600 bp mRNA linear EST 21-OCT-2003
DEFINITION	854999 MARC 4PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION	CF787130
VERSION	CF787130.1 GI:37791691
KEYWORDS	EST.
SOURCE	Sus scrofa (pig)
ORGANISM	Sus scrofa Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE	1 (bases 1 to 600)
AUTHORS	Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE	Porcine EST collection using a normalized library constructed from embryos representing early developmental stages
JOURNAL	Unpublished (2003)
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329. Plate: TMW8001 row: I column: 11 Seq primer: GTAATACGACTCACTATAGGG.

FEATURES	Location/Qualifiers
source	1. .600 /organism="Sus scrofa" /mol_type="mRNA" /db_xref="taxon:9823" /tissue_type="pooled" /lab_host="DH10B" /clone_lib="MARC 4PIG" /note="Vector: pcDNA3.1; Site_1: EcoRI; Site_2: NotI; Library made with combined RNA from day-10, day-13, day-15, day-25, and day-30 whole embryos."

ORIGIN

Alignment Scores:

Pred. No.:	21.9	Length:	600
Score:	94.50	Matches:	32
Percent Similarity:	44.83%	Conservative:	20
Best Local Similarity:	27.59%	Mismatches:	41
Query Match:	14.52%	Indels:	23
DB:	14	Gaps:	6

US-09-684-215B-23 (1-128) x CF787130 (1-600)

```
Qy          12 GlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArg 31
               |||     |||     |||     |||                 ||||||:::|||
Db         554 GGGGCAACAGGACAAGGCAGGCCGGCTTCTCAGGGGTCAGGGGCATTTGGGCAGATGCGC 495

Qy          32 SerGlyGlyGlySerProThrVal---HisIleGlyProThrAlaPheLeuGlyLeuGly 50
                     |||||:::|||:::|||     :::   |||   |||     |||   |||
Db         494 TTGAGTGGGGGGGCGCCCTCCGTGTCCTCTGTATCACCCTGGGCTGCCTCGGGGACAGGT 435

Qy          51 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70
                        |||   |||::::::                 |||::::::   :::
Db         434 --------------GGCACTGGCTCAAAGACGGGGTAAGCTTCAGGGGCCTCAGCCTCC 390

Qy          71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 90
               ::::::   |||     ||||||     :::   ::::::   |||:::   :::
Db         389 AGCAATTCTGGA------ACAGGCTCCTCCTTGGTCAGCATCGGGGGTTCATCCGTGCCG 336

Qy          91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
                  |||                     |||     ||||||         |||:::
Db         335 CCTGCA----------------------GGTATGCCCCCAGGA---------TCTCTC 309

Qy         111 ThrTrpGlnThr-----------LysSerGlyGlyThrArgThrGly 122
               :::|||   :::           |||   |||||   :::|||
Db         308 TCCTGGCCCAGCTGTGGGGGCTGAGCATCCTCAGGGACCCCTTCCGGT 261
```

RESULT 26
BX674871/c

LOCUS	BX674871 647 bp mRNA linear EST 28-OCT-2003
DEFINITION	BX674871 Sus Scrofa library (scac) Sus scrofa cDNA clone scac0037i.j.05 5prim, mRNA sequence.
ACCESSION	BX674871
VERSION	BX674871.1 GI:38008823
KEYWORDS	EST.
SOURCE	Sus scrofa (pig)
ORGANISM	Sus scrofa Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE	1 (bases 1 to 647)
AUTHORS	Bonnet,A., Tosser-Klopp,G., Benne,F., Cabau,C., Villeger,S., Soares,M., Bonaldo,F. and Hatey,F.
TITLE	A Pig Normalised Multi-Tissue cDNA Library
JOURNAL	Unpublished (2003)
COMMENT	Contact: Tosser-Klopp G Genetique Animale Institut National de la Recherche Agronomique Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan cedex, FRANCE Tel: 33 (0) 5.61.28.51.14 Fax: 33 (0) 5.61.28.53.08 Email: tosser@toulouse.inra.fr Clone distribution: AGENAE Resource centre. Francois PIUMI, Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73 Sequence cleaned of vector, adaptator and repetitions. Contact us at sigenasupport@jouy.inra.fr to obtain the chromatogram of this sequence. Plate: 0037, row: j column: 5.

FEATURES	Location/Qualifiers
source	1. .647 /organism="Sus scrofa" /mol_type="mRNA" /db_xref="taxon:9823" /clone="scac0037i.j.05" /tissue_type="mixed" /clone_lib="Sus Scrofa library (scac)" /note="Vector: pT7T3D-pac vector; tissues: adipose tissue, brain, kidney, liver, muscle, ovary, testis, heart, hypothalamus, pancreas, skin, spleen, thymus, placenta, pituitary gland, seminal vesicle, small intestine, uterus, adrenals, bulbo uretral gland, cerebral trunk,

epididymis, female gonad, gall-bladder, hippocampus,
large intestine, male gonad, melanocytes, stomach, udder"
ORIGIN

Alignment Scores:

Pred. No.:	24.4	Length:	647
Score:	94.50	Matches:	32
Percent Similarity:	44.83%	Conservative:	20
Best Local Similarity:	27.59%	Mismatches:	41
Query Match:	14.52%	Indels:	23
DB:	13	Gaps:	6

US-09-684-215B-23 (1-128) x BX674871 (1-647)

```
Qy         12 GlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArg 31
              |||      |||         |||      |||                 ||||||:::|||
Db        600 GGGGCAACAGGACAAGGCAGGCCGGCTTCTCAGGGGTCAGGGGCATTTGGGCAGATGCGC 541

Qy         32 SerGlyGlyGlySerProThrVal---HisIleGlyProThrAlaPheLeuGlyLeuGly 50
                 ||||||:::|||:::|||      :::   |||  |||      |||   |||
Db        540 TTGAGTGGGGGGGCGCCCTCCGTGTCCTCTGTATCACCCTGGGCTGCCTCGGGGACAGGT 481

Qy         51 ValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGlySerAlaProAla 70
                       |||   |||::::::                 |||::::::   :::
Db        480 --------------GGCACTGGCTCAAAGACGGGGTAAGCTTCAGGGGCCTCAGCCTCC 436

Qy         71 AlaSerLeuGlyIleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsn 90
              ::::::   |||      ||||||      :::   ::::::   |||:::   :::
Db        435 AGCAATTCTGGA------ACAGGCTCCTCCTTGGTCAGCATCGGGGGTTCATCCGTGCCG 382

Qy         91 SerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerVal 110
                 |||                      |||      ||||||         |||:::
Db        381 CCTGCA----------------------GGTATGCCCCCAGGA---------TCTCTC 355

Qy        111 ThrTrpGlnThr------------LysSerGlyGlyThrArgThrGly 122
              :::|||   :::            |||   ||||||   :::|||
Db        354 TCCTGGCCCAGCTGTGGGGGCTGAGCATCCTCAGGGACCCCTTCCGGT 307
```

RESULT 27
BJ285991/c

```
LOCUS       BJ285991                728 bp    mRNA    linear   EST 09-APR-2002
DEFINITION  BJ285991 Y. Ogihara unpublished cDNA library, Wh_r Triticum
            aestivum cDNA clone whr13l24 3', mRNA sequence.
ACCESSION   BJ285991
VERSION     BJ285991.1  GI:20105650
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 728)
  AUTHORS   Ogihara,Y. and Murai,K.
  TITLE     Expressed genes in Triticum aestivum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .728
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="whr13l24"
                     /tissue_type="root"
                     /dev_stage="Feekes' scale 1"
                     /clone_lib="Y. Ogihara unpublished cDNA library, Wh_r"
ORIGIN
```

Alignment Scores:

Pred. No.:	29	Length:	728
Score:	94.50	Matches:	37
Percent Similarity:	42.22%	Conservative:	20
Best Local Similarity:	27.41%	Mismatches:	53
Query Match:	14.52%	Indels:	25
DB:	12	Gaps:	5

US-09-684-215B-23 (1-128) x BJ285991 (1-728)

```
Qy         10 SerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle--------- 26
              :::      |||   |||||||||||                :::   |||
Db        582 ACGTCCGCCGGTGTTGGTTTTGCTATCCAATCATCAACTATACTCAAAATAGTTCCTCAG 523

Qy         27 ---------AlaGlyGlnIleArgSerGlyGlySerProThrValHisIleGlyPro 43
                      :::|||::::::|||      |||  :::                 |||
Db        522 TTGATTCAGTCTGGAAAGGTTCGACGTGCTGGGCTGAATGTGGAATTCGCACCAGATCCG 463

Qy         44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63
              |||:::      |||  |||  |||            ||||||   :::   :::
Db        462 ATTGCGTAT---CAGCTTAACGTCCGTGAT-----------GGTGCTCTTATACTGAAG 418

Qy         64 ValValGlySerAlaProAlaAlaSerLeuGly-------------------------- 74
              |||         :::          |||
Db        417 GTTCCCCCGGGCAGCACCGTGGAAAAAGCAGGTCTAGTTCCTACAGGCAGGGGTTTTGCC 358

Qy         75 ------IleSerThrGlyAspValIleThrAlaValAspGlyAlaProIleAsnSerAla 92
                    |||      ||||||||||||   ||||||||||||   ||||||
Db        357 GGCAATATTGTCCTGGGCGACGTCATTGTTGCAGTGGATGGCAAACCGATTAAGGGCAAA 298

Qy         93 ThrAlaMetAlaAspAlaLeuAsnGlyHisHisProGlyAspValIleSerValThrTrp 112
              :::  :::         |||:::   :::      ||||||   :::::::::|||
Db        297 TCTGACCTCCTGAGGGTTCTAGACGACTACGGGGTCGGAGACACGGTGACCCTGACGATC 238

Qy        113 GlnThrLysSerGlyGlyThrArgThrGlyAsnValThrLeuAla 127
              :::         ||||||   |||      :::    |||:::
Db        237 AGAAGA---------GGCACTGAGACAATTCCGATAGCCCTGTCG 202
```

RESULT 28
CC345457

```
LOCUS       CC345457                600 bp    DNA     linear   GSS 16-MAY-2003
DEFINITION  OGQAF70TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0365K20,
            genomic survey sequence.
ACCESSION   CC345457
VERSION     CC345457.1  GI:30814864
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE     Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBMa0365K20"
```

/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"
ORIGIN

Alignment Scores:
Pred. No.: 27.5 Length: 600
Score: 93.50 Matches: 32
Percent Similarity: 41.58% Conservative: 10
Best Local Similarity: 31.68% Mismatches: 48
Query Match: 14.36% Indels: 11
DB: 28 Gaps: 2

US-09-684-215B-23 (1-128) x CC345457 (1-600)

```
Qy          13 GlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIleArgSer 32
               |||   |||||||||||||||     |||     |||      ::::::
Db         303 GGCATTGGTTTTGCCATTCCGACCGCGCTGGCGACCAAAATTATGGATAAGCTGATTCGC 362

Qy          33 GlyGlyGlySerProThrValHisIleGlyProThrAlaPheLeuGlyLeuGlyValVal 52
                 |||                 :::||||||   |||         |||       :::
Db         363 GATGGCCGGGTGATCCGCGGTTATATCGGTATTACC---------GGGCGCGAGCTGCCG 413

Qy          53 AspAsnAsnGlyAsnGlyAlaArgValGlnArgValValGly------------------ 66
                  :::|||   |||:::   :::   ||||||   |||
Db         414 CCGCTGCACGGTCAGGGCAGCAACCTCGATCGGGTGCAGGGCATTATTGTTAGTGAAGTG 473

Qy          67 ------SerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 84
                   ||||||         ||||||         ||||||:::      |||
Db         474 ACGCCGGGCGGTCCGGCGGATCGCGCCGGCATTCAGGCCAATGATGTCCTGCTGCAGGTA 533

Qy          85 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 104
               :::|||   |||      ||||||           |||   :::          |||
Db         534 AACGGTAAGCCGGCGCTGTCCGCGCAGGAGACCATGGACCAGGTGGCGGAAATCCGCCCG 593

Qy         105 Gly 105
               |||
Db         594 GGA 596
```

RESULT 29
CA471201/c
LOCUS CA471201 907 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10698245 NCI_CGAP_ZKid1 Danio rerio cDNA clone IMAGE:6791062 5', mRNA sequence.
ACCESSION CA471201
VERSION CA471201.1 GI:24927553
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 907)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14294 row: b column: 21
High quality sequence stop: 590.
FEATURES Location/Qualifiers
source 1. .907
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6791062"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_ZKid1"
/note="Organ: kidney; Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.: 50.1 Length: 907
Score: 93.50 Matches: 32
Percent Similarity: 41.18% Conservative: 17
Best Local Similarity: 26.89% Mismatches: 53
Query Match: 14.36% Indels: 17
DB: 14 Gaps: 4

US-09-684-215B-23 (1-128) x CA471201 (1-907)

```
Qy          12 GlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle---------AlaGly 28
               |||             :::::::|||   |||   |||:::::::::
Db         847 GGGCACTTGAACCTTTCATTGCCCAATGGTGATGCCCTGTCCGTCAGGCAGCTGGTAGCT 788

Qy          29 GlnIleArgSerGlyGlyGlySerProThrValHisIle---------------GlyPro 43
                 ::::::   ||||||||||||         |||:::               |||
Db         787 CTTCTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCATCTCCTGTTCAAAGTCAAGGGCCAC 728

Qy          44 ThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArg 63
                  |||   |||   |||   |||   ||||||
Db         727 ATAGCAGAGCTTCTCCTTGATGTCACGGACAAT-----------------TTCCCTCTC 686

Qy          64 ValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAla 83
                  |||                 ||||||||||||           :::|||:::::
Db         685 AGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAGGATCTTCATGAGGTAGTCAGTCAG 626

Qy          84 ValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla---------LeuAsn 100
               |||      :::   :::   :::        |||                  ::::::
Db         625 GTCACGGCCAGCCAAGTCCAGACGGAGGATGGCATGGGGCAGGGCGTAACCCTCGTAGAT 566

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThr 119
               ||||||      ||||||   |||:::                ||||||   |||
Db         565 GGGCACAGTGTGGGTGACACCATCACCAGAGTCCATCACGATACCAGTGGTACGACC 509
```

RESULT 30
CF265501/c
LOCUS CF265501 913 bp mRNA linear EST 12-AUG-2003
DEFINITION AGENCOURT_15130858 NIH_ZGC_7 Danio rerio cDNA clone IMAGE:6996778 5', mRNA sequence.
ACCESSION CF265501
VERSION CF265501.1 GI:33605880
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 913)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```
           Plate: LLAM14682  row: j  column: 09
           High quality sequence stop: 620.
FEATURES             Location/Qualifiers
     source          1. .913
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:6996778"
                     /tissue_type="whole body"
                     /lab_host="DH10B"
                     /clone_lib="NIH_ZGC_7"
                     /note="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
                     Bulk tissue was collected from a whole adult individual
                     from the Tuebingen strain. 1st strand cDNA was primed with
                     a Not I - oligo(dT) primer, double-stranded cDNA was
                     cloned into the Not I and EcoRV sites of pExpress-1.
                     Library was size-selected for >1 kb fragments and
                     normalized. A non-normalized version of this library is
                     also available (NIH_ZGC_10). Library was constructed by
                     Open Biosystems (Huntsville, AL)"
ORIGIN

Alignment Scores:
Pred. No.:                50.6          Length:          913
Score:                    93.50         Matches:         37
Percent Similarity:       37.96%        Conservative:    15
Best Local Similarity:    27.01%        Mismatches:      56
Query Match:              14.36%        Indels:          29
DB:                       14            Gaps:            5

US-09-684-215B-23 (1-128) x CF265501 (1-913)

Qy          11 GlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||   :::|||      |||:::       :::   :::       :::
Db         826 CAGGGCCCTCGGGGGCACCATGAACCTCTTCATGGCCAATTGGTGATGACCTGTTCCGTC 767

Qy          31 Arg--------------------SerGlyGlyGlySerProThrValHisIle------ 41
               |||                    ||||||||||||        |||:::
Db         766 AGGCAGTTCGTAGCTCTTCTCCAGGGAGGAGGAGGAAGCAGCGGTGCCCATCTCCTGCTC 707

Qy          42 ---------GlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGly 58
                      |||      |||   |||   |||   |||   ||||||
Db         706 AAAGTCAAGGGCCACATAGCAGAGCTTCTCCTTGATGTCACGGACAAT------------ 659

Qy          59 AlaArgValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGly 78
                              |||             ||||||||||||
Db         658 ------TTCCCTCTCAGCTGTGGTGGTGAAGCTGTAGCCTCTCTCGGTCAGGATCTTCAT 605

Qy          79 AspValIleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAla 98
               :::|||::::::   |||     :::   :::   :::       |||
Db         604 GAGGTAGTCAGTCAGGTCACGGCCAGCCAAGTCCAGACGGAGGATGGCATGGGGCAGGGC 545

Qy          99 ---------LeuAsnGlyHisHisProGlyAspValIleSerValThrTrpGlnThrLys 115
                      ::::::||||||      ||||||   |||:::
Db         544 GTAACCCTCGTAGATGGGCACAGTGTGGGTGACACCATCACCAGAGTCCATCACGATACC 485

Qy         116 SerGlyGlyThrArg-----------------------ThrGlyAsnVal 124
               ||||||   ||||||                      |||||||||
Db         484 AGTGGTACGACCAGAGGCATACAGGGACAGCACAGCCTGGATGGCAACGTA 434


Search completed: April 30, 2004, 04:41:02
Job time : 2510.02 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:28:57 ; Search time 65.5311 Seconds
(without alignments)
1083.969 Million cell updates/sec

Title: US-09-684-215B-23
Perfect score: 651
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....SVTWQTKSGGTRTGNVTLAE 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13249/app_query.fasta_1.1180
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_115_@runat_29042004_061306_13249 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	651	100.0	702	4	US-09-287-849-27	Sequence 27, Appl
2	651	100.0	2287	4	US-09-223-040-1	Sequence 1, Appli
3	651	100.0	2287	4	US-09-287-849-1	Sequence 1, Appli
4	651	100.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
5	651	100.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
6	646	99.2	447	3	US-08-818-112-4	Sequence 4, Appli
7	646	99.2	447	4	US-08-818-111-4	Sequence 4, Appli
8	646	99.2	447	4	US-09-056-556-4	Sequence 4, Appli
9	646	99.2	447	4	US-09-072-596-4	Sequence 4, Appli
10	646	99.2	447	4	US-09-072-967-4	Sequence 4, Appli
11	646	99.2	1872	3	US-08-818-112-17	Sequence 17, Appl
12	646	99.2	1872	4	US-08-818-111-17	Sequence 17, Appl
13	646	99.2	1872	4	US-09-056-556-17	Sequence 17, Appl
14	646	99.2	1872	4	US-09-072-596-17	Sequence 17, Appl
15	646	99.2	1872	4	US-09-072-967-17	Sequence 17, Appl
16	608	93.4	675	4	US-09-636-215-822	Sequence 822, App
17	608	93.4	675	4	US-09-685-166A-822	Sequence 822, App
18	608	93.4	822	4	US-09-736-457-1862	Sequence 1862, Ap
19	608	93.4	900	4	US-09-643-597-353	Sequence 353, App
20	608	93.4	900	4	US-09-606-421B-353	Sequence 353, App
21	608	93.4	915	4	US-09-636-215-834	Sequence 834, App
22	608	93.4	915	4	US-09-685-166A-834	Sequence 834, App
23	608	93.4	945	4	US-09-736-457-1861	Sequence 1861, Ap
24	608	93.4	1012	4	US-09-643-597-351	Sequence 351, App
25	608	93.4	1012	4	US-09-606-421B-351	Sequence 351, App
26	608	93.4	1203	4	US-09-636-215-851	Sequence 851, App
27	608	93.4	1203	4	US-09-685-166A-851	Sequence 851, App
28	608	93.4	1464	4	US-09-620-412C-348	Sequence 348, App
29	608	93.4	1464	4	US-09-598-419-348	Sequence 348, App
30	608	93.4	1557	4	US-09-620-412C-332	Sequence 332, App
31	608	93.4	1557	4	US-09-598-419-332	Sequence 332, App
32	608	93.4	1578	4	US-09-556-877-188	Sequence 188, App
33	608	93.4	1578	4	US-09-620-412C-188	Sequence 188, App
34	608	93.4	1578	4	US-09-598-419-188	Sequence 188, App
35	608	93.4	1752	4	US-09-620-412C-352	Sequence 352, App
36	608	93.4	1752	4	US-09-598-419-352	Sequence 352, App
37	608	93.4	1758	4	US-09-620-412C-336	Sequence 336, App
38	608	93.4	1758	4	US-09-598-419-336	Sequence 336, App
39	608	93.4	1860	4	US-09-620-412C-308	Sequence 308, App
40	608	93.4	1860	4	US-09-598-419-308	Sequence 308, App
41	608	93.4	1896	4	US-09-620-412C-324	Sequence 324, App
42	608	93.4	1896	4	US-09-598-419-324	Sequence 324, App
43	608	93.4	1941	4	US-09-620-412C-316	Sequence 316, App
44	608	93.4	1941	4	US-09-598-419-316	Sequence 316, App
45	608	93.4	1965	4	US-09-620-412C-340	Sequence 340, App

ALIGNMENTS

RESULT 1
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 702
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
; OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
; OTHER INFORMATION: frame 1

```
;   NAME/KEY: CDS
;   LOCATION: (1)..(693)
;   OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
;   OTHER INFORMATION: reading frame 1
;   NAME/KEY: CDS
;   LOCATION: (2)..(700)
;   OTHER INFORMATION: reading frame 2
;   NAME/KEY: CDS
;   LOCATION: (3)..(701)
;   OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:
Pred. No.:              4.03e-64        Length:         702
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-287-849-27 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAG 408

RESULT 2
US-09-223-040-1
; Sequence 1, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION:  Fusion Proteins of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE:  014058-009010US
;  CURRENT APPLICATION NUMBER: US/09/223,040
;  CURRENT FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ra12-TbH9-Ra35
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-09-223-040-1

Alignment Scores:
Pred. No.:              1.92e-63        Length:         2287
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-223-040-1 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 3
US-09-287-849-1
; Sequence 1, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION:  Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE:  014058-009020US
;  CURRENT APPLICATION NUMBER: US/09/287,849
;  CURRENT FILING DATE:  1999-04-07
;  PRIOR APPLICATION NUMBER:  US 08/818,112
;  PRIOR FILING DATE:  1997-03-13
;  PRIOR APPLICATION NUMBER:  US 08/942,578
;  PRIOR FILING DATE:  1997-10-01
;  PRIOR APPLICATION NUMBER:  US 09/025,197
;  PRIOR FILING DATE:  1998-02-18
```

```
;  PRIOR APPLICATION NUMBER:  US 09/056,556
;  PRIOR FILING DATE:  1998-04-07
;  PRIOR APPLICATION NUMBER:  US 09/223,040
;  PRIOR FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  46
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2287
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
;   OTHER INFORMATION: protein Ra12-TbH9-Ra35 (designated Mtb32-Mtb39
;   OTHER INFORMATION: fusion)
;   NAME/KEY: modified_base
;   LOCATION: (30)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: modified_base
;   LOCATION: (33)
;   OTHER INFORMATION: n = g, a, c or t
;   NAME/KEY: CDS
;   LOCATION: (42)..(2231)
;   NAME/KEY: modified_base
;   LOCATION: (2270)
;   OTHER INFORMATION: n = g, a, c or t
US-09-287-849-1

Alignment Scores:
Pred. No.:                1.92e-63        Length:         2287
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-287-849-1 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAG 446

RESULT 4
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.:                4.17e-59        Length:         4403765
Score:                    651.00          Matches:        128
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       3               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-103-840A-2 (1-4403765)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151984 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 152043

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152044 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 152103

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152104 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 152163

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152164 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152223

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152224 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152283

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152284 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152343

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db     152344 ACAGGGAACGTGACATTGGCCGAG 152367

RESULT 5
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
```

```
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:              4.18e-59        Length:         4411529
Score:                  651.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     3               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-103-840A-1 (1-4411529)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151815 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 151874

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151875 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 151934

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151935 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 151994

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151995 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152054

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152055 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152114

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152115 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152174

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db     152175 ACAGGGAACGTGACATTGGCCGAG 152198

RESULT 6
US-08-818-112-4
; Sequence 4, Application US/08818112
; Patent No. 6290969
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  153
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,112
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-112-4

Alignment Scores:
Pred. No.:              8.06e-64        Length:         447
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     3               Gaps:           0

US-09-684-215B-23 (1-128) x US-08-818-112-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 7
US-08-818-111-4
; Sequence 4, Application US/08818111
; Patent No. 6338852
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
```

```
;       STATE:  Washington
;       COUNTRY:  USA
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/818,111
;       FILING DATE:  13-MAR-1997
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maki, David J.
;       REGISTRATION NUMBER:  31,392
;       REFERENCE/DOCKET NUMBER:  210121.417C6
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  447 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-818-111-4

Alignment Scores:
Pred. No.:              8.06e-64        Length:         447
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-23 (1-128) x US-08-818-111-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 8
US-09-056-556-4
; Sequence 4, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND TREATM
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-4

Alignment Scores:
Pred. No.:              8.06e-64        Length:         447
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-056-556-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 9
US-09-072-596-4
```

```
; Sequence 4, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-4

Alignment Scores:
Pred. No.:                8.06e-64        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-072-596-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 10
US-09-072-967-4
; Sequence 4, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,967
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-967-4

Alignment Scores:
Pred. No.:                8.06e-64        Length:         447
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-072-967-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db         311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
                ||||||||||||||||||||||||
Db         371 ACAGGGAACGTGACATTGGCCGAG 394

RESULT 11
US-08-818-112-17
; Sequence 17, Application US/08818112
; Patent No. 6290969
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  153
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,112
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-112-17

Alignment Scores:
Pred. No.:              5.34e-63        Length:         1872
Score:                  646.00          Matches:        127
Percent Similarity:     99.22%          Conservative:   0
Best Local Similarity:  99.22%          Mismatches:     1
Query Match:            99.23%          Indels:         0
DB:                     3               Gaps:           0

US-09-684-215B-23 (1-128) x US-08-818-112-17 (1-1872)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy         121 ThrGlyAsnValThrLeuAlaGlu 128
                ||||||||||||||||||||||||
Db        1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 12
US-08-818-111-17
; Sequence 17, Application US/08818111
; Patent No. 6338852
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,111
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  1872 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-818-111-17

Alignment Scores:
Pred. No.:                5.34e-63        Length:         1872
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-23 (1-128) x US-08-818-111-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 13
US-09-056-556-17
; Sequence 17, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND          TREATM
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-17

Alignment Scores:
Pred. No.:                5.34e-63        Length:         1872
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-056-556-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 14
US-09-072-596-17
; Sequence 17, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-17

Alignment Scores:
Pred. No.:                5.34e-63        Length:         1872
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-072-596-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141

RESULT 15
US-09-072-967-17
; Sequence 17, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,967
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1872 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-967-17

Alignment Scores:
Pred. No.:                5.34e-63        Length:         1872
Score:                    646.00          Matches:        127
Percent Similarity:       99.22%          Conservative:   0
Best Local Similarity:    99.22%          Mismatches:     1
Query Match:              99.23%          Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-23 (1-128) x US-09-072-967-17 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
```

```
Db       1118 ACAGGGAACGTGACATTGGCCGAG 1141
```

RESULT 16
US-09-636-215-822
; Sequence 822, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-822

Alignment Scores:

Pred. No.:	2.48e-59	Length:	675
Score:	608.00	Matches:	122
Percent Similarity:	96.09%	Conservative:	1
Best Local Similarity:	95.31%	Mismatches:	1
Query Match:	93.39%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-23 (1-128) x US-09-636-215-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
```

```
Db       370 ACAGGGAACGTGACATTGGCCGAG 393
```

RESULT 17
US-09-685-166A-822
; Sequence 822, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-822

Alignment Scores:

Pred. No.:	2.48e-59	Length:	675
Score:	608.00	Matches:	122
Percent Similarity:	96.09%	Conservative:	1
Best Local Similarity:	95.31%	Mismatches:	1
Query Match:	93.39%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-23 (1-128) x US-09-685-166A-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
```

```
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 18
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1862

Alignment Scores:
Pred. No.:                3.22e-59        Length:         822
Score:                    608.00          Matches:        122
Percent Similarity:       96.09%          Conservative:   1
Best Local Similarity:    95.31%          Mismatches:     1
Query Match:              93.39%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-23 (1-128) x US-09-736-457-1862 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 19
US-09-643-597-353
; Sequence 353, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-353

Alignment Scores:
Pred. No.:                3.63e-59        Length:         900
Score:                    608.00          Matches:        122
Percent Similarity:       96.09%          Conservative:   1
Best Local Similarity:    95.31%          Mismatches:     1
Query Match:              93.39%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-23 (1-128) x US-09-643-597-353 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 20
US-09-606-421B-353
; Sequence 353, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
```

```
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-606-421B-353

Alignment Scores:
Pred. No.:                3.63e-59        Length:         900
Score:                    608.00          Matches:        122
Percent Similarity:       96.09%          Conservative:   1
Best Local Similarity:    95.31%          Mismatches:     1
Query Match:              93.39%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-23 (1-128) x US-09-606-421B-353 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 21
US-09-636-215-834
; Sequence 834, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
```

```
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 834
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-636-215-834

Alignment Scores:
Pred. No.:                3.71e-59        Length:         915
Score:                    608.00          Matches:        122
Percent Similarity:       96.09%          Conservative:   1
Best Local Similarity:    95.31%          Mismatches:     1
Query Match:              93.39%          Indels:         4
DB:                       4               Gaps:           1

US-09-684-215B-23 (1-128) x US-09-636-215-834 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::           ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              ||||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 22
US-09-685-166A-834
; Sequence 834, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
```

```
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-834

Alignment Scores:
Pred. No.:              3.71e-59      Length:         915
Score:                  608.00        Matches:        122
Percent Similarity:     96.09%        Conservative:   1
Best Local Similarity:  95.31%        Mismatches:     1
Query Match:            93.39%        Indels:         4
DB:                     4             Gaps:           1

US-09-684-215B-23 (1-128) x US-09-685-166A-834 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 23
US-09-736-457-1861
; Sequence 1861, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1861
;   LENGTH: 945
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1861

Alignment Scores:
Pred. No.:              3.87e-59      Length:         945
Score:                  608.00        Matches:        122
Percent Similarity:     96.09%        Conservative:   1
Best Local Similarity:  95.31%        Mismatches:     1
Query Match:            93.39%        Indels:         4
DB:                     4             Gaps:           1

US-09-684-215B-23 (1-128) x US-09-736-457-1861 (1-945)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 24
US-09-643-597-351
; Sequence 351, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 351
;   LENGTH: 1012
;   TYPE: DNA
```

; ORGANISM: Homo sapiens
US-09-643-597-351

Alignment Scores:
Pred. No.: 4.24e-59 Length: 1012
Score: 608.00 Matches: 122
Percent Similarity: 96.09% Conservative: 1
Best Local Similarity: 95.31% Mismatches: 1
Query Match: 93.39% Indels: 4
DB: 4 Gaps: 1

US-09-684-215B-23 (1-128) x US-09-643-597-351 (1-1012)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::              ||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 174

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 234

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 294

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 354

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 414

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        415 ACAGGGAACGTGACATTGGCCGAG 438
```

RESULT 25
US-09-606-421B-351
; Sequence 351, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 351
; LENGTH: 1012
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-351

Alignment Scores:
Pred. No.: 4.24e-59 Length: 1012
Score: 608.00 Matches: 122
Percent Similarity: 96.09% Conservative: 1
Best Local Similarity: 95.31% Mismatches: 1
Query Match: 93.39% Indels: 4
DB: 4 Gaps: 1

US-09-684-215B-23 (1-128) x US-09-606-421B-351 (1-1012)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::              ||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 174

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 234

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 294

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 354

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 414

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        415 ACAGGGAACGTGACATTGGCCGAG 438
```

RESULT 26
US-09-636-215-851
; Sequence 851, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-851

Alignment Scores:
Pred. No.: 5.32e-59 Length: 1203
Score: 608.00 Matches: 122
Percent Similarity: 96.09% Conservative: 1
Best Local Similarity: 95.31% Mismatches: 1

Query Match: 93.39% Indels: 4
DB: 4 Gaps: 1

US-09-684-215B-23 (1-128) x US-09-636-215-851 (1-1203)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||::::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393
```

RESULT 27
US-09-685-166A-851
; Sequence 851, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-851

Alignment Scores:
Pred. No.: 5.32e-59 Length: 1203
Score: 608.00 Matches: 122
Percent Similarity: 96.09% Conservative: 1
Best Local Similarity: 95.31% Mismatches: 1

Query Match: 93.39% Indels: 4
DB: 4 Gaps: 1

US-09-684-215B-23 (1-128) x US-09-685-166A-851 (1-1203)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||::::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393
```

RESULT 28
US-09-620-412C-348
; Sequence 348, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-348

Alignment Scores:
Pred. No.: 6.9e-59 Length: 1464
Score: 608.00 Matches: 122
Percent Similarity: 96.09% Conservative: 1
Best Local Similarity: 95.31% Mismatches: 1
Query Match: 93.39% Indels: 4
DB: 4 Gaps: 1

US-09-684-215B-23 (1-128) x US-09-620-412C-348 (1-1464)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||::::                ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG-----------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189
```

```
Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393
```

RESULT 29
US-09-598-419-348
; Sequence 348, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-348

Alignment Scores:

Pred. No.:	6.9e-59	Length:	1464
Score:	608.00	Matches:	122
Percent Similarity:	96.09%	Conservative:	1
Best Local Similarity:	95.31%	Mismatches:	1
Query Match:	93.39%	Indels:	4
DB:	4	Gaps:	1

US-09-684-215B-23 (1-128) x US-09-598-419-348 (1-1464)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::              ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393

RESULT 30
US-09-620-412C-332
; Sequence 332, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 332
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Chlymadia trachomatis
US-09-620-412C-332

Alignment Scores:

| | | | |
|---|---|---|---|
| Pred. No.: | 7.48e-59 | Length: | 1557 |
| Score: | 608.00 | Matches: | 122 |
| Percent Similarity: | 96.09% | Conservative: | 1 |
| Best Local Similarity: | 95.31% | Mismatches: | 1 |
| Query Match: | 93.39% | Indels: | 4 |
| DB: | 4 | Gaps: | 1 |

US-09-684-215B-23 (1-128) x US-09-620-412C-332 (1-1557)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||:::              ||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAG------------CTTCCCACCGTTCAT 129

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 189

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 249

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 309

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 369

Qy        121 ThrGlyAsnValThrLeuAlaGlu 128
              |||||||||||||||||||||||
Db        370 ACAGGGAACGTGACATTGGCCGAG 393
```

Search completed: April 30, 2004, 05:12:07
Job time : 713.531 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:28:57 ; Search time 65.5311 Seconds
(without alignments)
1083.969 Million cell updates/sec

Title: US-09-684-215B-18
Perfect score: 653
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13249/app_query.fasta_1.1180
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_115_@runat_29042004_061306_13249 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	653	100.0	675	4	US-09-636-215-822	Sequence 822, App
2	653	100.0	675	4	US-09-685-166A-822	Sequence 822, App
3	653	100.0	822	4	US-09-736-457-1862	Sequence 1862, Ap
4	653	100.0	900	4	US-09-643-597-353	Sequence 353, App
5	653	100.0	900	4	US-09-606-421B-353	Sequence 353, App
6	653	100.0	915	4	US-09-636-215-834	Sequence 834, App
7	653	100.0	915	4	US-09-685-166A-834	Sequence 834, App
8	653	100.0	945	4	US-09-736-457-1861	Sequence 1861, Ap
9	653	100.0	1012	4	US-09-643-597-351	Sequence 351, App
10	653	100.0	1012	4	US-09-606-421B-351	Sequence 351, App
11	653	100.0	1203	4	US-09-636-215-851	Sequence 851, App
12	653	100.0	1203	4	US-09-685-166A-851	Sequence 851, App
13	653	100.0	1464	4	US-09-620-412C-348	Sequence 348, App
14	653	100.0	1464	4	US-09-598-419-348	Sequence 348, App
15	653	100.0	1557	4	US-09-620-412C-332	Sequence 332, App
16	653	100.0	1557	4	US-09-598-419-332	Sequence 332, App
17	653	100.0	1578	4	US-09-556-877-188	Sequence 188, App
18	653	100.0	1578	4	US-09-620-412C-188	Sequence 188, App
19	653	100.0	1578	4	US-09-598-419-188	Sequence 188, App
20	653	100.0	1752	4	US-09-620-412C-352	Sequence 352, App
21	653	100.0	1752	4	US-09-598-419-352	Sequence 352, App
22	653	100.0	1758	4	US-09-620-412C-336	Sequence 336, App
23	653	100.0	1758	4	US-09-598-419-336	Sequence 336, App
24	653	100.0	1860	4	US-09-620-412C-308	Sequence 308, App
25	653	100.0	1860	4	US-09-598-419-308	Sequence 308, App
26	653	100.0	1896	4	US-09-620-412C-324	Sequence 324, App
27	653	100.0	1896	4	US-09-598-419-324	Sequence 324, App
28	653	100.0	1941	4	US-09-620-412C-316	Sequence 316, App
29	653	100.0	1941	4	US-09-598-419-316	Sequence 316, App
30	653	100.0	1965	4	US-09-620-412C-340	Sequence 340, App
31	653	100.0	1965	4	US-09-598-419-340	Sequence 340, App
32	653	100.0	2052	4	US-09-620-412C-356	Sequence 356, App
33	653	100.0	2052	4	US-09-598-419-356	Sequence 356, App
34	653	100.0	2076	4	US-09-620-412C-312	Sequence 312, App
35	653	100.0	2076	4	US-09-598-419-312	Sequence 312, App
36	653	100.0	2103	4	US-09-620-412C-344	Sequence 344, App
37	653	100.0	2103	4	US-09-598-419-344	Sequence 344, App
38	653	100.0	2148	4	US-09-620-412C-320	Sequence 320, App
39	653	100.0	2148	4	US-09-620-412C-328	Sequence 328, App
40	653	100.0	2148	4	US-09-598-419-320	Sequence 320, App
41	653	100.0	2148	4	US-09-598-419-328	Sequence 328, App
42	632	96.8	702	4	US-09-287-849-27	Sequence 27, Appl
43	632	96.8	2287	4	US-09-223-040-1	Sequence 1, Appli
44	632	96.8	2287	4	US-09-287-849-1	Sequence 1, Appli
45	632	96.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-636-215-822
; Sequence 822, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-822

Alignment Scores:

```
Pred. No.:              1.06e-67        Length:         675
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-636-215-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 2
US-09-685-166A-822
; Sequence 822, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-822

Alignment Scores:
```

```
Pred. No.:              1.06e-67        Length:         675
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-685-166A-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 3
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1862

Alignment Scores:
Pred. No.:              1.39e-67        Length:         822
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0
```

US-09-684-215B-18 (1-128) x US-09-736-457-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 4
US-09-643-597-353
; Sequence 353, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-353

Alignment Scores:

Pred. No.:	1.58e-67	Length:	900
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-643-597-353 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 5
US-09-606-421B-353
; Sequence 353, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-353

Alignment Scores:

Pred. No.:	1.58e-67	Length:	900
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-606-421B-353 (1-900)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

```
Qy          61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy          81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy         101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy         121 ThrLeuAlaGluGlyProProAla 128
               |||||||||||||||||||||||
Db         382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 6
US-09-636-215-834
; Sequence 834, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-834

Alignment Scores:

Pred. No.:	1.61e-67	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-636-215-834 (1-915)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy          41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201
```

```
Qy          61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy          81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy         101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy         121 ThrLeuAlaGluGlyProProAla 128
               |||||||||||||||||||||||
Db         382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 7
US-09-685-166A-834
; Sequence 834, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-834

Alignment Scores:

Pred. No.:	1.61e-67	Length:	915
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-685-166A-834 (1-915)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy          41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201
```

```
Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 8
US-09-736-457-1861
; Sequence 1861, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1861
;   LENGTH: 945
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1861

Alignment Scores:
Pred. No.:              1.69e-67      Length:         945
Score:                  653.00        Matches:        128
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     4             Gaps:           0

US-09-684-215B-18 (1-128) x US-09-736-457-1861 (1-945)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 9
US-09-643-597-351
; Sequence 351, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 351
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-351

Alignment Scores:
Pred. No.:              1.85e-67      Length:         1012
Score:                  653.00        Matches:        128
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     4             Gaps:           0

US-09-684-215B-18 (1-128) x US-09-643-597-351 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 186

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 246

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 306

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 366

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 426

Qy        121 ThrLeuAlaGluGlyProProAla 128
```

```
                    ||||||||||||||||||||||||
Db        427 ACATTGGCCGAGGGACCCCCGGCC 450

RESULT 10
US-09-606-421B-351
; Sequence 351, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 351
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-606-421B-351

Alignment Scores:
Pred. No.:                1.85e-67      Length:          1012
Score:                    653.00        Matches:         128
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-684-215B-18 (1-128) x US-09-606-421B-351 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 186

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 246

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 306

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 366

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 426

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        427 ACATTGGCCGAGGGACCCCCGGCC 450

RESULT 11
US-09-636-215-851
; Sequence 851, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-636-215-851

Alignment Scores:
Pred. No.:                2.34e-67      Length:          1203
Score:                    653.00        Matches:         128
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-684-215B-18 (1-128) x US-09-636-215-851 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 12
US-09-685-166A-851
; Sequence 851, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
```

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-851

Alignment Scores:

Pred. No.:	2.34e-67	Length:	1203
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-685-166A-851 (1-1203)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 13
US-09-620-412C-348
; Sequence 348, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-348

Alignment Scores:

Pred. No.:	3.07e-67	Length:	1464
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-620-412C-348 (1-1464)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 14
US-09-598-419-348
; Sequence 348, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-348

Alignment Scores:

```
Pred. No.:              3.07e-67        Length:         1464
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-598-419-348 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 15
US-09-620-412C-332
; Sequence 332, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 332
;   LENGTH: 1557
;   TYPE: DNA
;   ORGANISM: Chlymadia trachomatis
US-09-620-412C-332

Alignment Scores:
Pred. No.:              3.34e-67        Length:         1557
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-620-412C-332 (1-1557)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 16
US-09-598-419-332
; Sequence 332, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 332
;   LENGTH: 1557
;   TYPE: DNA
;   ORGANISM: Chlymadia trachomatis
US-09-598-419-332

Alignment Scores:
Pred. No.:              3.34e-67        Length:         1557
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-598-419-332 (1-1557)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 17
US-09-556-877-188
; Sequence 188, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 188
;   LENGTH: 1578
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-556-877-188

Alignment Scores:
Pred. No.:                3.4e-67        Length:          1578
Score:                    653.00         Matches:         128
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       4              Gaps:            0

US-09-684-215B-18 (1-128) x US-09-556-877-188 (1-1578)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 18
US-09-620-412C-188
; Sequence 188, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 188
;   LENGTH: 1578
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-620-412C-188

Alignment Scores:
Pred. No.:                3.4e-67        Length:          1578
Score:                    653.00         Matches:         128
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       4              Gaps:            0

US-09-684-215B-18 (1-128) x US-09-620-412C-188 (1-1578)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 19
US-09-598-419-188
; Sequence 188, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 188
;   LENGTH: 1578
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-598-419-188

Alignment Scores:
Pred. No.:                3.4e-67        Length:          1578
```

```
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-598-419-188 (1-1578)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 20
US-09-620-412C-352
; Sequence 352, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 352
;   LENGTH: 1752
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-352

Alignment Scores:
Pred. No.:              3.92e-67        Length:         1752
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-620-412C-352 (1-1752)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 21
US-09-598-419-352
; Sequence 352, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 352
;   LENGTH: 1752
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-352

Alignment Scores:
Pred. No.:              3.92e-67        Length:         1752
Score:                  653.00          Matches:        128
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-18 (1-128) x US-09-598-419-352 (1-1752)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381
```

```
Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 22
US-09-620-412C-336
; Sequence 336, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 336
;   LENGTH: 1758
;   TYPE: DNA
;   ORGANISM: Chlymadia trachomatis
US-09-620-412C-336

Alignment Scores:

Pred. No.:	3.94e-67	Length:	1758
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-620-412C-336 (1-1758)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 23
US-09-598-419-336
; Sequence 336, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 336
;   LENGTH: 1758
;   TYPE: DNA
;   ORGANISM: Chlymadia trachomatis
US-09-598-419-336

Alignment Scores:

Pred. No.:	3.94e-67	Length:	1758
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-598-419-336 (1-1758)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 24
US-09-620-412C-308
; Sequence 308, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 308
;   LENGTH: 1860
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-308

Alignment Scores:

Pred. No.:	4.25e-67	Length:	1860
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-620-412C-308 (1-1860)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 25
US-09-598-419-308
; Sequence 308, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 308
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-308

Alignment Scores:

Pred. No.:	4.25e-67	Length:	1860
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-598-419-308 (1-1860)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 26
US-09-620-412C-324
; Sequence 324, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 324
; LENGTH: 1896
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-324

Alignment Scores:

Pred. No.:	4.37e-67	Length:	1896
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-620-412C-324 (1-1896)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              |||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405
```

```
RESULT 27
US-09-598-419-324
; Sequence 324, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 324
;   LENGTH: 1896
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-324

Alignment Scores:
Pred. No.:                4.37e-67      Length:          1896
Score:                    653.00        Matches:         128
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-684-215B-18 (1-128) x US-09-598-419-324 (1-1896)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 28
US-09-620-412C-316
; Sequence 316, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 316
;   LENGTH: 1941
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-316

Alignment Scores:
Pred. No.:                4.51e-67      Length:          1941
Score:                    653.00        Matches:         128
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-684-215B-18 (1-128) x US-09-620-412C-316 (1-1941)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy         41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy         61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy         81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy        101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy        121 ThrLeuAlaGluGlyProProAla 128
              ||||||||||||||||||||||||
Db        382 ACATTGGCCGAGGGACCCCCGGCC 405

RESULT 29
US-09-598-419-316
; Sequence 316, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 316
;   LENGTH: 1941
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-316

Alignment Scores:
Pred. No.:                4.51e-67      Length:          1941
Score:                    653.00        Matches:         128
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       4             Gaps:            0

US-09-684-215B-18 (1-128) x US-09-598-419-316 (1-1941)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
```

```
Db        22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy        21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40

Db        82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy        41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60

Db       142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy        61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80

Db       202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy        81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100

Db       262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy       101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120

Db       322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy       121 ThrLeuAlaGluGlyProProAla 128

Db       382 ACATTGGCCGAGGGACCCCCGGCC 405
```

RESULT 30
US-09-620-412C-340
; Sequence 340, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 340
; LENGTH: 1965
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-340

Alignment Scores:

Pred. No.:	4.59e-67	Length:	1965
Score:	653.00	Matches:	128
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-18 (1-128) x US-09-620-412C-340 (1-1965)

```
Qy         1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

Db        22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy        21 GlyGlnAlaMetAlaIleAlaGlyGlnIleLysLeuProThrValHisIleGlyProThr 40

Db        82 GGGCAGGCGATGGCGATCGCGGGCCAGATCAAGCTTCCCACCGTTCATATCGGGCCTACC 141

Qy        41 AlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArgValGlnArgVal 60

Db       142 GCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGAGTCCAACGCGTG 201

Qy        61 ValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspValIleThrAlaVal 80

Db       202 GTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACCGCGGTC 261

Qy        81 AspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsnGlyHisHisPro 100

Db       262 GACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAACGGGCATCATCCC 321

Qy       101 GlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArgThrGlyAsnVal 120

Db       322 GGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTG 381

Qy       121 ThrLeuAlaGluGlyProProAla 128

Db       382 ACATTGGCCGAGGGACCCCCGGCC 405
```

Search completed: April 30, 2004, 05:01:19
Job time : 76.5311 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2004, 06:21:52 ; Search time 22 Seconds
(without alignments)
300.369 Million cell updates/sec

Title: US-09-684-215B-18
Perfect score: 653
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	653	100.0	224	4	US-09-636-215-825	Sequence 825, App
2	653	100.0	224	4	US-09-685-166A-825	Sequence 825, App
3	653	100.0	267	4	US-09-643-597-352	Sequence 352, App
4	653	100.0	267	4	US-09-606-421B-352	Sequence 352, App
5	653	100.0	273	4	US-09-736-457-1864	Sequence 1864, Ap
6	653	100.0	299	4	US-09-643-597-354	Sequence 354, App
7	653	100.0	299	4	US-09-606-421B-354	Sequence 354, App
8	653	100.0	304	4	US-09-636-215-835	Sequence 835, App
9	653	100.0	304	4	US-09-685-166A-835	Sequence 835, App
10	653	100.0	314	4	US-09-736-457-1863	Sequence 1863, Ap
11	653	100.0	400	4	US-09-636-215-852	Sequence 852, App
12	653	100.0	400	4	US-09-685-166A-852	Sequence 852, App
13	653	100.0	487	4	US-09-620-412C-349	Sequence 349, App
14	653	100.0	487	4	US-09-598-419-349	Sequence 349, App
15	653	100.0	518	4	US-09-620-412C-333	Sequence 333, App
16	653	100.0	518	4	US-09-598-419-333	Sequence 333, App
17	653	100.0	525	4	US-09-556-877-196	Sequence 196, App
18	653	100.0	525	4	US-09-620-412C-196	Sequence 196, App
19	653	100.0	525	4	US-09-598-419-196	Sequence 196, App
20	653	100.0	583	4	US-09-620-412C-353	Sequence 353, App
21	653	100.0	583	4	US-09-598-419-353	Sequence 353, App
22	653	100.0	585	4	US-09-620-412C-337	Sequence 337, App
23	653	100.0	585	4	US-09-598-419-337	Sequence 337, App
24	653	100.0	619	4	US-09-620-412C-309	Sequence 309, App
25	653	100.0	619	4	US-09-598-419-309	Sequence 309, App
26	653	100.0	631	4	US-09-620-412C-325	Sequence 325, App
27	653	100.0	631	4	US-09-598-419-325	Sequence 325, App
28	653	100.0	646	4	US-09-620-412C-317	Sequence 317, App
29	653	100.0	646	4	US-09-598-419-317	Sequence 317, App
30	653	100.0	654	4	US-09-620-412C-341	Sequence 341, App
31	653	100.0	654	4	US-09-598-419-341	Sequence 341, App
32	653	100.0	683	4	US-09-620-412C-357	Sequence 357, App
33	653	100.0	683	4	US-09-598-419-357	Sequence 357, App
34	653	100.0	691	4	US-09-620-412C-313	Sequence 313, App
35	653	100.0	691	4	US-09-598-419-313	Sequence 313, App
36	653	100.0	700	4	US-09-620-412C-345	Sequence 345, App
37	653	100.0	700	4	US-09-598-419-345	Sequence 345, App
38	653	100.0	715	4	US-09-620-412C-321	Sequence 321, App
39	653	100.0	715	4	US-09-620-412C-329	Sequence 329, App
40	653	100.0	715	4	US-09-598-419-321	Sequence 321, App
41	653	100.0	715	4	US-09-598-419-329	Sequence 329, App
42	632	96.8	231	4	US-09-287-849-28	Sequence 28, Appl
43	632	96.8	729	4	US-09-223-040-2	Sequence 2, Appli
44	632	96.8	729	4	US-09-287-849-2	Sequence 2, Appli
45	627	96.0	132	3	US-08-818-112-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1
US-09-636-215-825
; Sequence 825, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 825
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-636-215-825

Query Match 100.0%; Score 653; DB 4; Length 224;
Best Local Similarity 100.0%; Pred. No. 1e-67;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135
```

RESULT 2
US-09-685-166A-825
; Sequence 825, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 825
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-825

Query Match 100.0%; Score 653; DB 4; Length 224;
Best Local Similarity 100.0%; Pred. No. 1e-67;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
Db 8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy 61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
Db 68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy 121 TLAEGPPA 128
Db 128 TLAEGPPA 135

RESULT 3
US-09-643-597-352
; Sequence 352, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 352
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-352

Query Match 100.0%; Score 653; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3e-67;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
Db 8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy 61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
Db 68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy 121 TLAEGPPA 128
Db 128 TLAEGPPA 135

RESULT 4
US-09-606-421B-352
; Sequence 352, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 352
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-352

Query Match 100.0%; Score 653; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3e-67;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
Db 8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy 61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
Db 68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy 121 TLAEGPPA 128
Db 128 TLAEGPPA 135

RESULT 5
US-09-736-457-1864

; Sequence 1864, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1864
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-1864

Query Match 100.0%; Score 653; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.4e-67;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135
```

RESULT 6
US-09-643-597-354
; Sequence 354, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 354
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-354

Query Match 100.0%; Score 653; DB 4; Length 299;
Best Local Similarity 100.0%; Pred. No. 1.5e-67;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135
```

RESULT 7
US-09-606-421B-354
; Sequence 354, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 354
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-354

Query Match 100.0%; Score 653; DB 4; Length 299;
Best Local Similarity 100.0%; Pred. No. 1.5e-67;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135
```

RESULT 8
US-09-636-215-835
; Sequence 835, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.

```
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 835
;   LENGTH: 304
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-636-215-835

  Query Match             100.0%;  Score 653;  DB 4;  Length 304;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 9
US-09-685-166A-835
; Sequence 835, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 835
;   LENGTH: 304
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-685-166A-835

  Query Match             100.0%;  Score 653;  DB 4;  Length 304;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 10
US-09-736-457-1863
; Sequence 1863, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1863
;   LENGTH: 314
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-736-457-1863

  Query Match             100.0%;  Score 653;  DB 4;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 11
US-09-636-215-852
; Sequence 852, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
```

```
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-636-215-852

  Query Match             100.0%;  Score 653;  DB 4;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 12
US-09-685-166A-852
; Sequence 852, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 852
;   LENGTH: 400
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-685-166A-852

  Query Match             100.0%;  Score 653;  DB 4;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 13
US-09-620-412C-349
; Sequence 349, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 349
;   LENGTH: 487
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-349

  Query Match             100.0%;  Score 653;  DB 4;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 14
US-09-598-419-349
; Sequence 349, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 349
;   LENGTH: 487
```

```
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-349

  Query Match            100.0%;  Score 653;  DB 4;  Length 487;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 15
US-09-620-412C-333
; Sequence 333, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 333
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Chlymadia trachomatis
US-09-620-412C-333

  Query Match            100.0%;  Score 653;  DB 4;  Length 518;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 16
US-09-598-419-333
; Sequence 333, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 333
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Chlymadia trachomatis
US-09-598-419-333

  Query Match            100.0%;  Score 653;  DB 4;  Length 518;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 17
US-09-556-877-196
; Sequence 196, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 196
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-556-877-196

  Query Match            100.0%;  Score 653;  DB 4;  Length 525;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 18
US-09-620-412C-196
; Sequence 196, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 196
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-620-412C-196

  Query Match             100.0%;  Score 653;  DB 4;  Length 525;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 19
US-09-598-419-196
; Sequence 196, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 196
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-598-419-196

  Query Match             100.0%;  Score 653;  DB 4;  Length 525;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 20
US-09-620-412C-353
; Sequence 353, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 353
;   LENGTH: 583
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-353

  Query Match             100.0%;  Score 653;  DB 4;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 21
US-09-598-419-353
; Sequence 353, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 353
;   LENGTH: 583
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-353

  Query Match             100.0%;  Score 653;  DB 4;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 22
US-09-620-412C-337
; Sequence 337, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 337
;   LENGTH: 585
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-337

  Query Match             100.0%;  Score 653;  DB 4;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 4e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 23
US-09-598-419-337
; Sequence 337, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 337
;   LENGTH: 585
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-337

  Query Match             100.0%;  Score 653;  DB 4;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 4e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 24
US-09-620-412C-309
; Sequence 309, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 309
;   LENGTH: 619
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-309

  Query Match             100.0%;  Score 653;  DB 4;  Length 619;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 25
US-09-598-419-309
; Sequence 309, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 309
;   LENGTH: 619
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-309

  Query Match             100.0%;  Score 653;  DB 4;  Length 619;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 26
US-09-620-412C-325
; Sequence 325, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 325
;   LENGTH: 631
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-325

  Query Match             100.0%;  Score 653;  DB 4;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 27
US-09-598-419-325
; Sequence 325, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 325
;   LENGTH: 631
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-325

  Query Match             100.0%;  Score 653;  DB 4;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 28
US-09-620-412C-317
; Sequence 317, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 317
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-317

  Query Match             100.0%;  Score 653;  DB 4;  Length 646;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 29
US-09-598-419-317
; Sequence 317, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 317
;   LENGTH: 646
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-317

  Query Match             100.0%;  Score 653;  DB 4;  Length 646;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 30
US-09-620-412C-341
; Sequence 341, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 341
;   LENGTH: 654
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-341

  Query Match             100.0%;  Score 653;  DB 4;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 31
US-09-598-419-341
; Sequence 341, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 341
;   LENGTH: 654
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-341

  Query Match             100.0%;  Score 653;  DB 4;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 32
US-09-620-412C-357
; Sequence 357, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 357
;   LENGTH: 683
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-357

  Query Match             100.0%;  Score 653;  DB 4;  Length 683;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 33
US-09-598-419-357
; Sequence 357, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 357
;   LENGTH: 683
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-357

  Query Match             100.0%;  Score 653;  DB 4;  Length 683;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 34
US-09-620-412C-313
; Sequence 313, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 313
;   LENGTH: 691
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-313

  Query Match             100.0%;  Score 653;  DB 4;  Length 691;
  Best Local Similarity   100.0%;  Pred. No. 5e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 35
US-09-598-419-313
; Sequence 313, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 313
;   LENGTH: 691
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-313

  Query Match             100.0%;  Score 653;  DB 4;  Length 691;
  Best Local Similarity   100.0%;  Pred. No. 5e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 36
US-09-620-412C-345
; Sequence 345, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 345
;   LENGTH: 700
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-345

  Query Match             100.0%;  Score 653;  DB 4;  Length 700;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 37
US-09-598-419-345
; Sequence 345, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 345
;   LENGTH: 700
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-345

  Query Match             100.0%;  Score 653;  DB 4;  Length 700;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 38
US-09-620-412C-321
; Sequence 321, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
```

```
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 321
;   LENGTH: 715
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-321

  Query Match             100.0%;  Score 653;  DB 4;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 39
US-09-620-412C-329
; Sequence 329, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 329
;   LENGTH: 715
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-329

  Query Match             100.0%;  Score 653;  DB 4;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 40
US-09-598-419-321
; Sequence 321, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 321
;   LENGTH: 715
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-321

  Query Match             100.0%;  Score 653;  DB 4;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


RESULT 41
US-09-598-419-329
; Sequence 329, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 329
;   LENGTH: 715
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-329

  Query Match             100.0%;  Score 653;  DB 4;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-67;
  Matches  128;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 TAASDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRV 67

Qy         61 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 VGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNV 127

Qy        121 TLAEGPPA 128
              ||||||||
Db        128 TLAEGPPA 135


Search completed: April 20, 2004, 06:25:00
Job time : 23 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:32:52 ; Search time 3784.74 Seconds
(without alignments)
1511.672 Million cell updates/sec

Title: US-09-684-215B-4
Perfect score: 675
Sequence: 1 TAASDNFQLSQGGQGFAIPI.....QTKSGGTRTGNVTLAEGPPA 132

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13185/app_query.fasta_1.1180
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_6034_@runat_29042004_061304_13185 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	675	100.0	702	6	BD251334	BD251334 Fused pro
2	675	100.0	702	6	AR403747	AR403747 Sequence
3	675	100.0	1068	6	BD274032	BD274032 Sequences
4	675	100.0	1068	6	AX005788	AX005788 Sequence
5	675	100.0	1143	6	BD274033	BD274033 Sequences
6	675	100.0	1143	6	AX005790	AX005790 Sequence
7	675	100.0	2287	6	BD251322	BD251322 Fused pro
8	675	100.0	2287	6	AR303127	AR303127 Sequence
9	675	100.0	2287	6	AR403735	AR403735 Sequence
10	675	100.0	14029	1	AE006925	AE006925 Mycobacte
11	675	100.0	341957	15	BX842572	Bx842572 Mycobacte
12	675	100.0	343050	1	BX248334	BX248334 Mycobacte
13	670	99.3	447	6	AR169152	AR169152 Sequence
14	670	99.3	447	6	AR182442	AR182442 Sequence
15	670	99.3	447	6	AR194825	AR194825 Sequence
16	670	99.3	447	6	AR233097	AR233097 Sequence
17	670	99.3	447	6	AR353302	AR353302 Sequence
18	670	99.3	447	6	AX429596	AX429596 Sequence
19	670	99.3	447	6	AX832581	AX832581 Sequence
20	670	99.3	447	6	BD006325	BD006325 Compounds
21	670	99.3	447	6	BD006445	BD006445 Compounds
22	670	99.3	447	6	BD069285	BD069285 Compounds
23	670	99.3	447	6	BD205817	BD205817 Compounds
24	670	99.3	1872	6	AR169165	AR169165 Sequence
25	670	99.3	1872	6	AR182455	AR182455 Sequence
26	670	99.3	1872	6	AR194838	AR194838 Sequence
27	670	99.3	1872	6	AR233110	AR233110 Sequence
28	670	99.3	1872	6	AR353315	AR353315 Sequence
29	670	99.3	1872	6	AX429609	AX429609 Sequence
30	670	99.3	1872	6	AX832594	AX832594 Sequence
31	670	99.3	1872	6	BD006338	BD006338 Compounds
32	670	99.3	1872	6	BD006458	BD006458 Compounds
33	670	99.3	1872	6	BD069298	BD069298 Compounds
34	670	99.3	1872	6	BD205830	BD205830 Compounds
35	632	93.6	675	6	AR261272	AR261272 Sequence
36	632	93.6	675	6	AR400535	AR400535 Sequence
37	632	93.6	675	6	AR405802	AR405802 Sequence
38	632	93.6	675	6	AX201049	AX201049 Sequence
39	632	93.6	675	6	AX267848	AX267848 Sequence
40	632	93.6	822	6	AR277645	AR277645 Sequence
41	632	93.6	822	6	AX369152	AX369152 Sequence
42	632	93.6	894	6	AX351489	AX351489 Sequence
43	632	93.6	900	6	AR220690	AR220690 Sequence
44	632	93.6	900	6	AX365960	AX365960 Sequence
45	632	93.6	915	6	AR261281	AR261281 Sequence

ALIGNMENTS

RESULT 1

```
BD251334
LOCUS       BD251334                 702 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof.
ACCESSION   BD251334
VERSION     BD251334.1  GI:33061104
KEYWORDS    JP 2002510494-A/13.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Skeiky,Y.A.W., Alderson,M. and Neto,A.C.
  TITLE     Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof
  JOURNAL   Patent: JP 2002510494-A 13 09-APR-2002;
            CORIXA CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002510494-A/13
            PD   09-APR-2002
            PF   07-APR-1999 JP 2000542460
            PR   07-APR-1998 US    09/056556,30-DEC-1998 US    09/223040 PI
            YASIR A W SKEIKY,MARK ALDERSON,ANTONIO CAMPOS NETO PC
            C12N15/09,A61K39/04,A61K48/00,A61P31/04,C07K14/35,C07K19/00, PC
            C12P21/02,
            PC   C12N15/00
            CC   Description of Artificial Sequence:bi-fusion
            protein Ra12-DPPD
            CC            (designated
            CC   Mtb24), reading frame 1
            CC   bi-fusion protein Ra12-DPPD (designated Mtb24),reading frame 1
            CC   reading frame 2
            CC   reading frame 3
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(696)
            FT   CDS             (2). .(700)
            FT   CDS             (3). .(701).
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:                2.37e-42        Length:          702
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-4 (1-132) x BD251334 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 2
AR403747
LOCUS       AR403747                 702 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 27 from patent US 6627198.
ACCESSION   AR403747
VERSION     AR403747.1  GI:40151423
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of Mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6627198-A 27 30-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                2.37e-42        Length:          702
Score:                    675.00          Matches:         132
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-4 (1-132) x AR403747 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 144

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 204

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 264

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 324

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 384

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||
Db        385 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 420

RESULT 3
BD274032
LOCUS       BD274032                 1068 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose.
```

```
ACCESSION   BD274032
VERSION     BD274032.1  GI:33083800
KEYWORDS    JP 2002534956-A/256.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 1068)
  AUTHORS   Gicquel,B., Portnoi,D., Lim,E., Pelicic,V., Guigueno,A. and
            Salmoniere,Y.G.D.L.
  TITLE     Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose
  JOURNAL   Patent: JP 2002534956-A 256 22-OCT-2002;
            INSTITUT PASTEUR
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002534956-A/256
            PD   22-OCT-2002
            PF   14-AUG-1998 JP 2000509849
            PR   14-AUG-1997 FR     97/10404,11-SEP-1997 FR     97/11325 PI
            BRIGITTE GICQUEL,DENIS PORTNOI,ENG-MONG LIM,VLADIMIR PELICIC, PI
            AGNES GUIGUENO,YVES GOGUET DE LA SALMONIERE
            PC   C12N15/09,A61K39/04,A61K48/00,A61P11/00,A61P31/04,C07K14/35,
            PC   C07K16/12,
            PC   C07K19/00,C12N1/21,C12P21/02,C12Q1/02,C12Q1/68,G01N33/50, PC
            G01N33/53//
            PC   (C12N15/09,C12R1:32),C12N15/00,(C12N15/00,C12R1:32) CC   50D
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(1065).
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1773"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-42        Length:         1068
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x BD274032 (1-1068)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065

RESULT 4
AX005788
LOCUS       AX005788                1068 bp    DNA     linear   PAT 24-AUG-2000
DEFINITION  Sequence 907 from Patent WO9909186.
ACCESSION   AX005788
VERSION     AX005788.1  GI:9928795
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Portnoi,D. and Guigueno,A.
  TITLE     Polypeptide nucleic sequences exported from mycobacteria, vectors
            comprising same and uses for diagnosing and preventing tuberculosis
  JOURNAL   Patent: WO 9909186-A 907 25-FEB-1999;
            PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES             Location/Qualifiers
     source          1. .1068
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1773"
     CDS             1. .1068
                     /note="unnamed protein product; 50D"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC05178.1"
                     /db_xref="GI:9928796"
                     /db_xref="REMTREMBL:CAC05178"
                     /translation="MSNSRRRSLRWSWLLSVLAAVGLGLATAPAQAAPPALSQDRFAD
                     FPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGAGTGIVIDPNGVVLTNNHVIAGATD
                     INAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSG
                     GQGGTPRAVPGRVVALGQTVQASDSLTGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQ
                     VVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVVD
                     NNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVT
                     WQTKSGGTRTGNVTLAEGPPA"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-42        Length:         1068
Score:                    675.00          Matches:        132
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x AX005788 (1-1068)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 729

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 789

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 849

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 909

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 969

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        970 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1029

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||||
Db       1030 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1065

RESULT 5
BD274033
LOCUS       BD274033               1143 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose.
ACCESSION   BD274033
VERSION     BD274033.1  GI:33083801
KEYWORDS    JP 2002534956-A/257.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 1143)
  AUTHORS   Gicquel,B., Portnoi,D., Lim,E., Pelicic,V., Guigueno,A. and
            Salmoniere,Y.G.D.L.
  TITLE     Sequences nuclieic acid from polypeptides exportes from
            mycobacteries, vector the comprenant and applications at diagnostic
            and the prevention from the tuberculose
  JOURNAL   Patent: JP 2002534956-A 257 22-OCT-2002;
            INSTITUT PASTEUR
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002534956-A/257
            PD   22-OCT-2002
            PF   14-AUG-1998 JP 2000509849
            PR   14-AUG-1997 FR     97/10404,11-SEP-1997 FR     97/11325 PI
            BRIGITTE GICQUEL,DENIS PORTNOI,ENG-MONG LIM,VLADIMIR PELICIC, PI
            AGNES GUIGUENO,YVES GOGUET DE LA SALMONIERE
            PC   C12N15/09,A61K39/04,A61K48/00,A61P11/00,A61P31/04,C07K14/35,
            PC   C07K16/12,
            PC   C07K19/00,C12N1/21,C12P21/02,C12Q1/02,C12Q1/68,G01N33/50, PC
            G01N33/53//
            PC   (C12N15/09,C12R1:32),C12N15/00,(C12N15/00,C12R1:32) CC   SEQ
            ID No. 50F
            FH   Key             Location/Qualifiers
            FT   CDS             (4). .(1140).
FEATURES              Location/Qualifiers
     source           1. .1143
                      /organism="Mycobacterium tuberculosis"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:1773"
ORIGIN

Alignment Scores:
Pred. No.:              3.85e-42        Length:         1143
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-4 (1-132) x BD274033 (1-1143)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1140

RESULT 6
AX005790
LOCUS       AX005790               1143 bp    DNA     linear   PAT 24-AUG-2000
DEFINITION  Sequence 909 from Patent WO9909186.
ACCESSION   AX005790
VERSION     AX005790.1  GI:9928797
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Portnoi,D. and Guigueno,A.
  TITLE     Polypeptide nucleic sequences exported from mycobacteria, vectors
            comprising same and uses for diagnosing and preventing tuberculosis
  JOURNAL   Patent: WO 9909186-A 909 25-FEB-1999;
            PORTNOI DENIS (FR); GUIGUENO AGNES (FR)
FEATURES              Location/Qualifiers
     source           1. .1143
                      /organism="Mycobacterium tuberculosis"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:1773"
     CDS              <4. .1143
                      /note="unnamed protein product; SEQ ID NO 50F"
                      /codon_start=1
                      /transl_table=11
                      /protein_id="CAC05179.1"
                      /db_xref="GI:9928798"
                      /db_xref="REMTREMBL:CAC05179"
                      /translation="KNPAARTLKAGTISDSYPDTGGYGMSNSRRRSLRWSWLLSVLAA
                      VGLGLATAPAQAAPPALSQDRFADFPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGA
                      GTGIVIDPNGVVLTNNHVIAGATDINAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAG
                      GLPSAAIGGGVAVGEPVVAMGNSGGQGGTPRAVPGRVVALGQTVQASDSLTGAEETLN
                      GLIQFDAAIQPGDSGGPVVNGLGQVVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQ
                      IRSGGGSPTVHIGPTAFLGLGVVDNNGNGARVQRVVGSAPAASLGISTGDVITAVDGA
                      PINSATAMADALNGHHPGDVISVTWQTKSGGTRTGNVTLAEGPPA"
ORIGIN

Alignment Scores:
Pred. No.:              3.85e-42        Length:         1143
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-4 (1-132) x AX005790 (1-1143)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 804

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 864

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 924

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 984

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1044

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 1104

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
                ||||||||||||||||||||||||||||||||||
Db       1105 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1140

RESULT 7
BD251322
LOCUS       BD251322                2287 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof.
ACCESSION   BD251322
VERSION     BD251322.1  GI:33061092
KEYWORDS    JP 2002510494-A/1.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Skeiky,Y.A.W., Alderson,M. and Neto,A.C.
  TITLE     Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof
  JOURNAL   Patent: JP 2002510494-A 1 09-APR-2002;
            CORIXA CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002510494-A/1
            PD   09-APR-2002
            PF   07-APR-1999 JP 2000542460
            PR   07-APR-1998 US    09/056556,30-DEC-1998 US    09/223040 PI
            YASIR A W SKEIKY,MARK ALDERSON,ANTONIO CAMPOS NETO PC
            C12N15/09,A61K39/04,A61K48/00,A61P31/04,C07K14/35,C07K19/00, PC
            C12P21/02,
            PC   C12N15/00
            CC   Description of Artificial Sequence:tri-fusion protein Ra12- CC
                 TbH9-Ra35
            CC   (designated Mtb32A)
            CC   n = g, a, c or t
            CC   n = g, a, c or t
            CC   n = g, a, c or t
            FH   Key             Location/Qualifiers
            FT   modified_base   (30)
            FT   modified_base   (33)
            FT   CDS             (42). .(2231)
            FT   modified_base   (2270).
FEATURES             Location/Qualifiers
     source          1. .2287
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:              7.67e-42        Length:         2287
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-4 (1-132) x BD251322 (1-2287)
```

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
                ||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458

RESULT 8
AR303127
LOCUS       AR303127                2287 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 1 from patent US 6544522.
ACCESSION   AR303127
VERSION     AR303127.1  GI:31691855
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Skeiky,Y., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6544522-A 1 08-APR-2003;
FEATURES             Location/Qualifiers
     source          1. .2287
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              7.67e-42        Length:         2287
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-4 (1-132) x AR303127 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

```
RESULT 9
AR403735
LOCUS       AR403735                2287 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1 from patent US 6627198.
ACCESSION   AR403735
VERSION     AR403735.1  GI:40151411
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2287)
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of Mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6627198-A 1 30-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .2287
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                  7.67e-42        Length:         2287
Score:                      675.00          Matches:        132
Percent Similarity:         100.00%         Conservative:   0
Best Local Similarity:      100.00%         Mismatches:     0
Query Match:                100.00%         Indels:         0
DB:                         6               Gaps:           0

US-09-684-215B-4 (1-132) x AR403735 (1-2287)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 122

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 182

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 242

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 302

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 362

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 422

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        423 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 458
```

```
RESULT 10
AE006925
LOCUS       AE006925                14029 bp    DNA     linear   BCT 27-APR-2001
DEFINITION  Mycobacterium tuberculosis CDC1551,  section 11 of 280 of the
            complete genome.
ACCESSION   AE006925 AE000516
VERSION     AE006925.1  GI:13879610
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis CDC1551
  ORGANISM  Mycobacterium tuberculosis CDC1551
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 14029)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Whole genome comparison of Mycobacterium tuberculosis clinical and
            laboratory strains
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 14029)
  AUTHORS   Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .14029
                     /organism="Mycobacterium tuberculosis CDC1551"
                     /mol_type="genomic DNA"
                     /strain="CDC1551"
                     /db_xref="taxon:83331"
                     /note="clinical strain"
     gene            complement(99. .533)
                     /gene="MT0129"
     CDS             complement(99. .533)
                     /gene="MT0129"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK44353.1"
                     /db_xref="GI:13879611"
                     /translation="MGEFDPKLRFAQSPVARLATSTPDGTPHLVPVVFALGARRPAEA
                     TGADVIYTAVDAKRKTTQRLRRLANLEHNPRASVLVDSYADDWTQLWWVRADGVAAIH
                     RDGEVMRAAYRLLRAKYTQYQSVPLNGPVIAIAVQRWASRHA"
     gene            complement(643. .1062)
                     /gene="MT0130"
     CDS             complement(643. .1062)
                     /gene="MT0130"
                     /note="identified by Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAK44354.1"
                     /db_xref="GI:13879612"
                     /translation="MASSSSLPLLRHCVSDDQVTVVGFDGDDLGKTARRIAALVVQRA
                     IFLNDRNTAVAHSGDDAVLGHAVLPGVPRDPDPLHASSMYSILGMCQSVNGRPFDAIA
                     LVSVRLCHVQTDPTDSCGGRDRPGQLPCAPLDYHRHH"
     gene            1047. .1415
                     /gene="MT0131"
     CDS             1047. .1415
                     /gene="MT0131"
                     /note="identified by match to  protein family HMM"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA-binding protein, CopG family"
                     /protein_id="AAK44355.1"
```

/db_xref="GI:13879613"
/translation="MTKKPRNPADYVIGDDVEVSDVDLKQEEVYVDGERLTDERVEQM
ASESLRLAREREANLIPGGKSLSGGSAHSPAVQVVVSKATHAKLKELARSRKMSVSKL
LRPVLDEFVQRETGRILPRR"

gene 1682. .3367
/gene="MT0132"

CDS 1682. .3367
/gene="MT0132"
/note="similar to GB:AL123456; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="PE_PGRS family protein"
/protein_id="AAK44356.1"
/db_xref="GI:13879614"
/translation="MFAGGGAGGLGRCVMSFVSVAPEIVVAAATDLAGIGSAISAANA
AAAAPTTAVLAAGADEVSAAIAALFSGHAQAYQALSAQAAAFHQQFVQTLAGGAGAYA
AAEAQVEQQLLAAINAPTQALLGRPLIGNGADGAPGTGQAGGAGGILYGNGGNGGSGA
AGQAGGAGGPAGLIGHGGSGGAGGSGAAGGAGGHGGWLWGNGGVGGSGGAGVGAGVAG
GHGGAGGAAGLWGAGGGGGNGGNGADANIVSGGDGGLGGAGGGGGWLYGDGGAGGHGG
QGAGGGAGGAGGDGGQGGAGRGLWGTGGAGGHGGQGGGTGGPPLPGQAGMGAAGGAGG
LIGNGGAGGDGGVGASGGVAGVXGAGGNAMLIGHGGAGGAGGDSSFANGAAGGAGGAG
GHLFGNGGSGGHGGAVTAGNTGIGGAGGVGGDARLIGHGGAGGAGGDRAGALVGRDGG
PGGNGGAGGQLYGNGGDGGPGGQGGQAFGANNIGGTGGAGGNGGPAILSGNGGNGGAG
GAGGAGGAGGGAGGVGGAGGAPGTGGTLQAAVSGLVTALFGAPGQPGDTGQPG"

gene 3519. .4586
/gene="MT0133"

CDS 3519. .4586
/gene="MT0133"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="serine protease, putative"
/protein_id="AAK44357.1"
/db_xref="GI:13879615"
/translation="MSNSRRRSLRWSWLLSVLAAVGLGLATAPAQAAPPALSQDRFAD
FPALPLDPSAMVAQVGPQVVNINTKLGYNNAVGAGTGIVIDPNGVVLTNNHVIAGATD
INAFSVGSGQTYGVDVVGYDRTQDVAVLQLRGAGGLPSAAIGGGVAVGEPVVAMGNSG
GQGGTPRAVPGRVVALGQTVQASDSLTGAEETLNGLIQFDAAIQPGDSGGPVVNGLGQ
VVGMNTAASDNFQLSQGGQGFAIPIGQAMAIAGQIRSGGGSPTVHIGPTAFLGLGVVD
NNGNGARVQRVVGSAPAASLGISTGDVITAVDGAPINSATAMADALNGHHPGDVISVT
WQTKSGGTRTGNVTLAEGPPA"

gene 4695. .6500
/gene="MT0134"

CDS 4695. .6500
/gene="MT0134"
/note="similar to GP:2808807; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="alpha-amylase family protein"
/protein_id="AAK44358.1"
/db_xref="GI:13879616"
/translation="MNEAEHSVEHPPVQGSHVEGGVVEHPDAKDFGSAAALPADPTWF
KHAVFYEVLVRAFFDASADGSGDLRGLIDRLDYLQWLGIDCIWLPPFYDSPLRDGGYD
IRDFYKVLPEFGTVDDFVALVDAAHRRGIRIITDLVMNHTSESHPWFQESRRDPDGPY
GDYYVWSDTSERYTDARIIFVDTEESNWSFDPVRRQFYWHRFFSHQPDLNYDNPAVQE
AMIDVIRFWLGLGIDGFRLDAVPYLFEREGTNCENLPETHAFLKRVRKVVDDEFPGRV
LLAEANQWPGDVVEYFGDPNTGGDECHMAFHFPLMPRIFMAVRRESRFPISEIIAQTP
PIPDMAQWGIFLRNHDELTLEMVTDEERDYMYAEYAKDPRMKANVGIRRRLAPLLDND
RNQIELFTALLLSLPGSPVLYYGDEIGMGDVIWLGDRDGVRIPMQWTPDRNAGFSTAN
PGRLYLPPSQDPVYGYQAVNVEAQRDTSTSLLNFTRTMLAVRRRHPAFAVGAFQELGG
SNPSVLAYVRQVAGDDGDTVLCVNNLSRFPQPIELDLQQWTNYTPVELTGHVEFPPRIG
QVPYLLTLPGHGFYWFQLTTHEVGAPPTCGGERRL"

gene 6603. .7970
/gene="MT0135"

CDS 6603. .7970
/gene="MT0135"
/note="similar to GP:2808801; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="pep2 protein"
/protein_id="AAK44359.1"
/db_xref="GI:13879617"
/translation="MTRSDTLATKLPWSDWLPRQRWYAGRNRELATVKPGVVVALRHN
LDLVLVDVTYTDGATERYQVLVGWDFEPASEYGTKAAIGVADDRTGFDALYDVAGPQF
LLSLIVSSAVCGTSTGEVTFTREPDVELPFAAQPRVCDAEQSNTSVIFDRRAILKVFR
RVSSGINPDIELNRVLTRAGNPHVARLLGAYQFGRPNRSPTDALAYALGMVTEYEANA
AEGWAMATASVRDLFAEGDLYAHEVGGDFAGESYRLGEAVASVHATLADSLGTAQATF
PVDRMLARLSSTVAVVPELREYAPTIEQQFQKLAAEAITVQRVHGDLHLGQVLRTPES
WLLIDFEGEPGQPLDERRAPDSPLRDVAGVLRSFEYAAYGPLVDQATDKQLAARAREW
VERNRAAFCDGYAVASGIDPRDSALLLGAYELDKAVYETGYETRHRPGWLPIPLRSIA
RLTAS"

gene 8038. .8817
/gene="MT0136"

CDS 8038. .8817
/gene="MT0136"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44360.1"
/db_xref="GI:13879618"
/translation="MQREIYDGEARLSWVLAALAGILGATAFTHSAGYFVTFMTGNSQ
RAVLGLFGDDAWMSVTASLLILFFVAGVVIASVCRRHFWAAHPHGPTVLTTFSLIFAA
GVDIMLGGWHESMLDFVPILFVVFGIGALNTSFVKDGEVSVPLSYVTGTLVKMGQGIE
RHLAGGKVEDWLGYFLLHASFVLGAAAGGAISMVVTGPQMLAVAAVVCAATTGYTYLH
ADRRGLVNQKRPQPGKRLFRALRRGELDSGTSTPATNYGSS"

gene complement(8949. .9989)
/gene="MT0137"

CDS complement(8949. .9989)
/gene="MT0137"
/note="similar to SP:P31953 GB:X57229 PID:48828
PID:1567734 PID:1877254; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="esterase, putative, antigen 85-C"
/protein_id="AAK44361.1"
/db_xref="GI:13879619"
/translation="MGLVVAMTFFEQVRRLRSAATTLPRRLAIAAMGAVLVYGLVGTF
GGPATAGAFSRPGLPVEYLQVPSASMGRDIKVQFQGGGPHAVYLLDGLRAQDDYNGWD
INTPAFEEYYQSGLSVIMPVGGQSSFYTDWYQPSQSNGQNYTYKWETFLTREMPAWLQ
ANKGVSPTGNAAVGLSMSGGSALILAAYYPQQFPYAASLSGFLNPSEGWWPTLIGLAM
NDSGGYNANSMWGPSSDPAWKRNDPMVQIPRLVANNTRIWVYCGNGTPSDLGGDNIPA
KFLEGLTLRTNQTFRDTYAADGGRNGVFNFPPNGTHSWPYWNEQLVAMKADIQHVLNG
ATPPAAPAAPAA"

gene 10218. .10673
/gene="MT0138"

CDS 10218. .10673
/gene="MT0138"
/note="similar to SP:P08634 PID:46222; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="nodulation protein N-related protein"
/protein_id="AAK44362.1"
/db_xref="GI:13879620"

Alignment Scores:

Pred. No.:	4.66e-41	Length:	14029
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-684-215B-4 (1-132) x AE006925 (1-14029)

Qy 1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

Db 4188 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 4247

Qy 21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

```
Db        4248 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 4307

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4308 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 4367

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4368 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 4427

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4428 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 4487

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4488 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 4547

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||||||
Db        4548 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 4583
```

```
RESULT 11
BX842572
ID   BX842572   standard; circular genomic DNA; PRO; 341957 BP.
XX
AC   BX842572; AL021427; AL021428; AL021926; AL021927; AL021928; AL021929;
AC   AL021930; AL123456; Z74410; Z80233; Z80775; Z86089; Z92669; Z92770; Z96071;
AC   Z97050;
XX
SV   BX842572.1
XX
DT   21-NOV-2003 (Rel. 77, Created)
DT   21-NOV-2003 (Rel. 77, Last updated, Version 1)
XX
DE   Mycobacterium tuberculosis H37Rv complete genome; segment 1/13
XX
KW   complete genome.
XX
OS   Mycobacterium tuberculosis H37Rv
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex; Mycobacterium tuberculosis.
XX
RN   [1]
RX   MEDLINE; 98295987.
RX   PUBMED; 9634230.
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry III C.E., Tekaia F., Badcock K.,
RA   Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K.,
RA   Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K.,
RA   Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J.,
RA   Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S.,
RA   Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the complete
RT   genome sequence";
RL   Nature 393:537-544(1998).
XX
RN   [2]
RX   PUBMED; 12368430.
RA   Camus J.C., Pryor M.J., Medigue C., Cole S.T.;
RT   "Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv";
RL   Microbiology 148:2967-2973(2002).
XX
RN   [3]
RP   1-341957
RA   Parkhill J.;
RT   ;
RL   Submitted (11-JUN-1998) to the EMBL/GenBank/DDBJ databases.
RL   Submitted on behalf of the Mycobacterium tuberculosis sequencing and
RL   mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton,
RL   Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut
RL   Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail:
RL   parkhill@sanger.ac.uk
XX
CC   Notes:
CC   Details of M. tuberculosis sequencing at the Sanger Centre
CC   are available on the World Wide Web.
CC   (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/)
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .341957
FT                   /db_xref="taxon:83332"
FT                   /mol_type="genomic DNA"
FT                   /organism="Mycobacterium tuberculosis H37Rv"
FT                   /strain="H37Rv"
FT   CDS             1. .1524
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0001, (MT0001, MTV029.01, P49993), len: 507 aa.
FT                   dnaA, chromosomal replication initiator protein (see
FT                   citations below), equivalent to other Mycobacterial
FT                   CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g.
FT                   P46388|DNAA_MYCLE from Mycobacterium leprae (502 aa);
FT                   Q9L7L7|DNAA_MYCPA from Mycobacterium paratuberculosis (509
FT                   aa); P49990|DNAA_MYCAV from Mycobacterium avium (508 aa);
FT                   P49992|DNAA_MYCSM from Mycobacterium smegmatis (504 aa);
FT                   etc. Also highly similar to others except in N-terminus
FT                   e.g. Q9ZH75|DNAA_STRCH CHROMOSOMAL REPLICATION INITIATOR
FT                   PROTEIN from Streptomyces chrysomallus (624 aa);
FT                   Q9ZH76|DNAA_STRRE from Streptomyces reticuli (643 aa);
FT                   DNAA_ECOLI|P03004|B3702 chromosomal replication initiator
FT                   protein from Escherichia coli strain K12 (467 aa), FASTA
FT                   scores: opt: 986, E(): 0, (43.2% identity in 389 aa
FT                   overlap); etc. Contains PS00017 ATP/GTP-binding site motif
FT                   A (P-loop) and PS01008 DnaA protein signature. BELONGS TO
FT                   THE DNAA FAMILY. Note that the first base of this gene has
FT                   been taken as base 1 of the Mycobacterium tuberculosis
FT                   H37Rv genomic sequence."
FT                   /transl_table=11
FT                   /function="PLAYS AN IMPORTANT ROLE IN THE INITIATION AND
FT                   REGULATION OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN
FT                   OF REPLICATION; IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA
FT                   AT A 9 BP CONSENSUS (DNAA BOX): 5'-TTATC(C/A)A(C/A)A-3'.
FT                   DNAA BINDS TO ATP AND TO ACIDIC PHOSPHOLIPIDS. DNAA PROTEIN
FT                   BINDS THE ORIGIN OF REPLICATION (oriC), ATP AND ADP, AND
FT                   EXHIBITED WEAK ATPase ACTIVITY."
FT                   /gene="dnaA"
FT                   /locus_tag="Rv0001"
FT                   /product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
FT                   /protein_id="CAA16238.1"
FT                   /translation="MTDDPGSGFTTVWNAVVSELNGDPKVDDGPSSDANLSAPLTPQQR
FT                   AWLNLVQPLTIVEGFALLSVPSSFVQNEIERHLRAPITDALSRRLGHQIQLGVRIAPPA
FT                   TDEADDTTVPPSENPATTSPDTTTDNDEIDDSAAARGDNQHSWPSYPTERPHNTDSATA
FT                   GVTSLNRRYTFDTFVIGASNRFAHAAALAIAEAPARAYNPLFIWGESGLGKTHLLHAAG
FT                   NYAQRLFPGMRVKYVSTEEFTNDFINSLRDDRKVAFKRSYRDVDVLLVDDIQFIEGKEG
FT                   IQEEFFHTFNTLHNANKQIVISSDRPPKQLATLEDRLRTRFEWGLITDVQPPELETRIA
FT                   ILRKKAQMERLAVPDDVLELIASSIERNIRELEGALIRVTAFASLNKTPIDKALAEIVL
FT                   RDLIADANTMQISAATIMAATAEYFDTTVEELRGPGKTRALAQSRQIAMYLCRELTDLS
FT                   LPKIGQAFGRDHTTVMYAQRKILSEMAERREVFDHVKELTTRIRQRSKR"
FT   misc_feature    622. .645
FT                   /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
FT   misc_feature    1384. .1440
FT                   /note="PS01008 DnaA protein signature"
FT   CDS             2052. .3260
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0002, (MTV029.02, MTCY10H4.0), len: 402 aa. dnaN,
FT                   DNA polymerase III (beta chain) (EC 2.7.7.7) (see citations
FT                   below), equivalent to other Mycobacterial DNA POLYMERASES
FT                   III BETA CHAIN e.g. NP_301130.1|NC_002677 from
FT                   Mycobacterium leprae (399 aa); Q9L7L6|DP3B_MYCPA from
FT                   Mycobacterium avium subsp. paratuberculosis (399 aa);
FT                   P52851|DP3B_MYCSM from Mycobacterium smegmatis (397 aa);
FT                   etc. Also highly similar to others e.g. P27903|DP3B_STRCO
FT                   DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor
FT                   (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity
```

```
FT                   in 337 aa overlap); P21174|DP3B_MICLU from Micrococcus
FT                   luteus (310 aa); P52023|DP3B_SYNP7 from Synechococcus sp.
FT                   strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in
FT                   neighbouring cosmid MTCY10H4.01."
FT                   /transl_table=11
FT                   /EC_number="2.7.7.7"
FT                   /function="DNA POLYMERASE III IS A COMPLEX, MULTICHAIN
FT                   ENZYME RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN
FT                   BACTERIA. THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5'
FT                   EXONUCLEASE ACTIVITY. THE BETA CHAIN IS REQUIRED FOR
FT                   INITIATION OF REPLICATION ONCE IT IS CLAMPED ONTO DNA, IT
FT                   SLIDES FREELY (BIDIRECTIONAL AND ATP-INDEPENDENT) ALONG
FT                   DUPLEX DNA [CATALYTIC ACTIVITY: N deoxynucleoside
FT                   triphosphate = N diphosphate + {DNA}N]."
FT                   /gene="dnaN"
FT                   /locus_tag="Rv0002"
FT                   /product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA
FT                   NUCLEOTIDYLTRANSFERASE)"
FT                   /protein_id="CAA16239.1"
FT                   /translation="MDAATTRVGLTDLTFRLLRESFADAVSWVAKNLPARPAVPVLSGV
FT                   LLTGSDNGLTISGFDYEVSAEAQVGAEIVSPGSVLVSGRLLSDITRALPNKPVDVHVEG
FT                   NRVALTCGNARFSLPTMPVEDYPTLPTLPEETGLLPAELFAEAISQVAIAAGRDDTLPM
FT                   LTGIRVEILGETVVLAATDRFRLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAGIGGS
FT                   DVRLSLGTGPGVGKDGLLGISGNGKRSTTRLLDAEFPKFRQLLPTEHTAVATMDVAELI
FT                   EAIKLVALVADRGAQVRMEFADGSVRLSAGADDVGRAEEDLVVDYAGEPLTIAFNPTYL
FT                   TDGLSSLRSERVSFGFTTAGKPALLRPVSGDDRPVAGLNGNGPFPAVSTDYVYLLMPVR
FT                   LPG"
FT   CDS             3280. .4437
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0003, (MTCY10H4.01), len: 385 aa. recF, DNA
FT                   replication and repair protein (see citations below),
FT                   equivalent to others Mycobacterial DNA replication and
FT                   repair proteins e.g. NP_301131.1|NC_002677 from
FT                   Mycobacterium leprae (385 aa); Q9L7L5|RECF_MYCPA from
FT                   Mycobacterium avium subsp. paratuberculosis (385 aa);
FT                   P50916|RECF_MYCSM from Mycobacterium smegmatis (384 aa);
FT                   etc. Also highly similar to others e.g. P36176|RECF_STRCO
FT                   DNA REPLICATION AND REPAIR PROTEIN from Streptomyces
FT                   coelicolor (373 aa); NP_440892.1|NC_000911 from
FT                   Synechocystis sp. strain PCC 6803 (384 aa);
FT                   NP_469352.1|NC_003212 from Listeria innocua (370 aa); etc.
FT                   Contains PS00017 ATP/GTP-binding site motif A (P-loop),
FT                   PS00617 RecF protein signature 1, and PS00618 RecF protein
FT                   signature 2. BELONGS TO THE RECF FAMILY."
FT                   /transl_table=11
FT                   /function="THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM
FT                   AND RECOMBINATION; IT IS REQUIRED FOR DNA REPLICATION AND
FT                   NORMAL SOS INDUCIBILITY. RECF BINDS PREFERENTIALLY TO
FT                   SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS TO BIND ATP."
FT                   /gene="recF"
FT                   /locus_tag="Rv0003"
FT                   /product="DNA REPLICATION AND REPAIR PROTEIN RECF
FT                   (SINGLE-STRAND DNA BINDING PROTEIN)"
FT                   /protein_id="CAB02424.1"
FT                   /translation="MYVRHLGLRDFRSWACVDLELHPGRTVFVGPNGYGKTNLIEALWY
FT                   STTLGSHRVSADLPLIRVGTDRAVISTIVVNDGRECAVDLEIATGRVNKARLNRSSVRS
FT                   TRDVVGVLRAVLFAPEDLGLVRGDPADRRRYLDDLAIVRRPAIAAVRAEYERVLRQRTA
FT                   LLKSVPGARYRGDRGVFDTLEVWDSRLAEHGAELVAARIDLVNQLAPEVKKAYQLLAPE
FT                   SRSASIGYRASMDVTGPSEQSDIDRQLLAARLLAALAARRDAELERGVCLVGPHRDDLI
FT                   LRLGDQPAKGFASHGEAWSLAVALRLAAYQLLRVDGGEPVLLLDDVFAELDVMRRRALA
FT                   TAAESAEQVLVTAAVLEDIPAGWDARRVHIDVRADDTGSMSVVLP"
FT   misc_feature    3367. .3390
FT                   /note="PS00017 ATP/GTP-binding site motif A"
FT   misc_feature    3634. .3690
FT                   /note="PS00617 RecF protein signature 1"
FT   misc_feature    4243. .4296
FT                   /note="PS00618 RecF protein signature 2"
FT   CDS             4434. .4997
FT                   /evidence=EXPERIMENTAL
FT                   /note="Rv0004, (MTCY10H4.02), len: 187 aa. Conserved
FT                   hypothetical protein (see Salazar et al., 1996), highly
FT                   similar, but longer 21 aa in N-terminus, to
FT                   AAF33696.1|AF222789 unknown protein from Mycobacterium
FT                   avium subsp. paratuberculosis (166 aa); and highly similar
FT                   to NP_301132.1|NC_002677 conserved hypothetical protein
FT                   from Mycobacterium leprae (189 aa); S70990 hypothetical
FT                   protein from Mycobacterium smegmatis (194 aa). Also highly
FT                   similar, except in N-terminal part, to

Alignment Scores:
Pred. No.:              1.12e-39        Length:         341957
Score:                  675.00          Matches:        132
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-4 (1-132) x BX842572 (1-341957)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151817 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 151876

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151877 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 151936

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151937 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 151996

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151997 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152056

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152057 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152116

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152117 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152176

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db     152177 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152212

RESULT 12
BX248334
LOCUS       BX248334               343050 bp    DNA     linear   BCT 11-JUN-2003
DEFINITION  Mycobacterium bovis subsp. bovis AF2122/97 complete genome; segment
            1/14.
ACCESSION   BX248334 BX248333
VERSION     BX248334.1  GI:31616762
KEYWORDS    complete genome.
SOURCE      Mycobacterium bovis subsp. bovis AF2122/97
  ORGANISM  Mycobacterium bovis subsp. bovis AF2122/97
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1
  AUTHORS   Garnier,T., Eiglmeier,K., Camus,J.-C., Medina,N., Mansoor,H.,
            Pryor,M., Duthoy,S., Grondin,S., Lacroix,C., Monsempe,C., Simon,S.,
            Harris,B., Atkin,R., Doggett,J., Mayes,R., Keating,L.,
            Wheeler,P.R., Parkhill,J., Barrell,B.G., Cole,S.T., Gordon,S.V. and
            Hewinson,G.
  TITLE     The complete genome sequence of Mycobacterium bovis
  JOURNAL   Online Publication
  REMARK    PNAS 10.1073/pnas.1130426100 ( Microbiology )
REFERENCE   2  (bases 1 to 343050)
  AUTHORS   Garnier,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-2003) Garnier T., Unite de Genetique Moleculaire
            Bacterienne Institut Pasteur 28,rue du Dr Roux 75724 PARIS cedex
            15, France. e-mail:tgarnier@pasteur.fr Submitted on behalf of the
            Mycobacterium bovis sequencing teams, TB Research Group, Veterinary
```

Laboratories Agency Weybridge, Woodham Lane, New Haw,Addlestone, Surrey KT15 3NB, UK. Sanger Centre, Wellcome Trust Genome Campus, Hinxton,Cambridge CB10 1SA, UK. PT4 Annotation, Genopole, Institut Pasteur, 28 Rue du Docteur Roux, 75724 Paris Cedex 15, France. Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

FEATURES Location/Qualifiers

source 1. .343050
/organism="Mycobacterium bovis subsp. bovis AF2122/97"
/mol_type="genomic DNA"
/strain="AF2122/97"
/db_xref="taxon:233413"

gene 1. .1524
/gene="dnaA"
/locus_tag="Mb0001"

CDS 1. .1524
/gene="dnaA"
/locus_tag="Mb0001"
/note="Mb0001, dnaA, len: 507 aa. Equivalent to Rv0001, len: 507 aa, from Mycobacterium tuberculosis strain H37Rv, (99.6% identity in 507 aa overlap). dnaA, chromosomal replication initiator protein (see citations below), equivalent to other Mycobacterial CHROMOSOMAL REPLICATION INITIATOR PROTEINS e.g. P46388|DNAA_MYCLE from Mycobacterium leprae (502 aa); Q9L7L7|DNAA_MYCPA from Mycobacterium paratuberculosis (509 aa); P49990|DNAA_MYCAV from Mycobacterium avium (508 aa); P49992|DNAA_MYCSM from Mycobacterium smegmatis (504 aa); etc. Also highly similar to others except in N-terminus e.g. Q9ZH75|DNAA_STRCH CHROMOSOMAL REPLICATION INITIATOR PROTEIN from Streptomyces chrysomallus (624 aa); Q9ZH76|DNAA_STRRE from Streptomyces reticuli (643 aa); DNAA_ECOLI|P03004|B3702 chromosomal replication initiator protein from Escherichia coli strain K12 (467 aa), FASTA scores: opt: 986, E(): 0, (43.2% identity in 389 aa overlap); etc. Contains PS00017 ATP/GTP-binding site motif A (P-loop) and PS01008 DnaA protein signature. BELONGS TO THE DNAA FAMILY. Note that the first base of this gene has been taken as base 1 of the Mycobacterium bovis genomic sequence."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="CHROMOSOMAL REPLICATION INITIATOR PROTEIN DNAA"
/protein_id="CAD92863.1"
/db_xref="GI:31616763"
/translation="MTDDPGSGFTTVWNAVVSELNGDPKVDDGPSSDANLSAPLTPQQ
RAWLNLVQPLTIVEGFALLSVPSSFVQNEIERHLRAPITDALSRRLGHQIQLGVRIAP
PATDEADDTTVPPSENPATTSPDTTTDNDEIDDSAAARGDNQHSWPSYFTERPRNTDS
ATAGVTSLNRRYTFDTFVIGASNRFAHAAALAIAEAPARAYNPLFIWGESGLGKTHLL
HAAGNYAQRLFPGMRVKYVSTEEFTNDFINSLRDDRKVAFKRSYRDVDVLLVDDIQFI
EGKEGIQEEFFHTFNTLHNANKQIVISSDRPPKQLATLEDRLRTRFEWGLITDVQPPE
LETRIAILRKKAQMERLAIPDDVLELIASSIERNIRELEGALIRVTAFASLNKTPIDK
ALAEIVLRDLIADANTMQISAATIMAATAEYFDTTVEELRGPGKTRALAQSRQIAMYL
CRELTDLSLPKIGQAFGRDHTTVMYAQRKILSEMAERREVFDHVKELTTRIRQRSKR"

gene 2052. .3260
/gene="dnaN"
/locus_tag="Mb0002"

CDS 2052. .3260
/gene="dnaN"
/locus_tag="Mb0002"
/EC_number="2.7.7.7"
/note="Mb0002, dnaN, len: 402 aa. Equivalent to Rv0002, len: 402 aa, from Mycobacterium tuberculosis strain H37Rv, (99.8% identity in 402 aa overlap). dnaN, DNA polymerase III (beta chain) (EC 2.7.7.7) (see citations below), equivalent to other Mycobacterial DNA POLYMERASES III BETA CHAIN e.g. NP_301130.1|NC_002677 from Mycobacterium leprae (399 aa); Q9L7L6|DP3B_MYCPA from Mycobacterium avium subsp. paratuberculosis (399 aa); P52851|DP3B_MYCSM from Mycobacterium smegmatis (397 aa); etc. Also highly similar to others e.g. P27903|DP3B_STRCO DNA POLYMERASE III BETA CHAIN from Streptomyces coelicolor (376 aa), FASTA scores: opt: 1189, E(): 0, (52.8% identity in 337 aa overlap); P21174|DP3B_MICLU from Micrococcus luteus (310 aa); P52023|DP3B_SYNP7 from Synechococcus sp. strain PCC 7942 (375 aa); etc. Overlaps and extends CDS in neighbouring cosmid MTCY10H4.01."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA POLYMERASE III (BETA CHAIN) DNAN (DNA NUCLEOTIDYLTRANSFERASE)"
/protein_id="CAD92864.1"
/db_xref="GI:31616764"
/translation="MDAATTRVGLTDLTFRLLRESFADAVSWVAKNLPARPAVPVLSG
VLLTGSDNGLTISGFDYEVSAEAQVGAEIVSPGSVLVSGRLLSDITRALPNKPVGVHV
EGNRVALTCGNARFSLPTMPVEDYPTLPTLPEETGLLPAELFAEAISQVAIAAGRDDT
LPMLTGIRVEILGETVVLAATDRFRLAVRELKWSASSPDIEAAVLVPAKTLAEAAKAG
IGGSDVRLSLGTGPGVGKDGLLGISGNGKRSTTRLLDAEFPKFRQLLPTEHTAVATMD
VAELIEAIKLVALVADRGAQVRMEFADGSVRLSAGADDVGRAEEDLVVDYAGEPLTIA
FNPTYLTDGLSSLRSERVSFGFTTAGKPALLRPVSGDDRPVAGLNGNGPFPAVSTDYV
YLLMPVRLPG"

gene 3280. .4437
/gene="recF"
/locus_tag="Mb0003"

CDS 3280. .4437
/gene="recF"
/locus_tag="Mb0003"
/note="Mb0003, recF, len: 385 aa. Equivalent to Rv0003, len: 385 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 385 aa overlap). recF, DNA replication and repair protein (see citations below), equivalent to others Mycobacterial DNA replication and repair proteins e.g. NP_301131.1|NC_002677 from Mycobacterium leprae (385 aa); Q9L7L5|RECF_MYCPA from Mycobacterium avium subsp. paratuberculosis (385 aa); P50916|RECF_MYCSM from Mycobacterium smegmatis (384 aa); etc. Also highly similar to others e.g. P36176|RECF_STRCO DNA REPLICATION AND REPAIR PROTEIN from Streptomyces coelicolor (373 aa); NP_440892.1|NC_000911 from Synechocystis sp. strain PCC 6803 (384 aa); NP_469352.1|NC_003212 from Listeria innocua (370 aa); etc. Contains PS00017 ATP/GTP-binding site motif A (P-loop), PS00617 RecF protein signature 1, and PS00618 RecF protein signature 2. BELONGS TO THE RECF FAMILY."
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA REPLICATION AND REPAIR PROTEIN RECF (SINGLE-STRAND DNA BINDING PROTEIN)"
/protein_id="CAD92865.1"
/db_xref="GI:31616765"
/translation="MYVRHLGLRDFRSWACVDLELHPGRTVFVGPNGYGKTNLIEALW
YSTTLGSHRVSADLPLIRVGTDRAVISTIVVNDGRECAVDLEIATGRVNKARLNRSSV
RSTRDVVGVLRAVLFAPEDLGLVRGDPADRRRYLDDLAIVRRPAIAAVRAEYERVVRQ
RTALLKSVPGARYRGDRGVFDTLEVWDSRLAEHGAELVAARIDLVNQLAPEVKKAYQL
LAPESRSASIGYRASMDVTGPSEQSDTDRQLLAARLLAALAARRDAELERGVCLVGPH
RDDLILRLGDQPAKGFASHGEAWSLAVALRLAAYQLLRVDGGEPVLLLDDVFAELDVM
RRRALATAAESAEQVLVTAAVLEDIPAGWDARRVHIDVRADDTGSMSVVLP"

gene 4434. .4997
/locus_tag="Mb0004"

CDS 4434. .4997
/locus_tag="Mb0004"
/note="Mb0004, -, len: 187 aa. Equivalent to Rv0004, len: 187 aa, from Mycobacterium tuberculosis strain H37Rv, (99.5% identity in 187 aa overlap). Conserved hypothetical protein (see citation below), highly similar, but longer 21 aa in N-terminus, to AAF33696.1|AF222789 unknown protein from Mycobacterium avium subsp. paratuberculosis (166 aa); and highly similar to NP_301132.1|NC_002677 conserved hypothetical protein from Mycobacterium leprae (189 aa); S70990 hypothetical protein from Mycobacterium smegmatis (194 aa). Also similar to in C-terminus to C-terminal part of P35925|YREG_STRCO HYPOTHETICAL 19.8 KDA PROTEIN (IN RECF-GYRB INTERGENIC REGION) from Streptomyces coelicolor (190 aa), FASTA scores: opt: 404, E(): 3.9e-18, (40.7% identity in 189 aa overlap)."

```
                     /codon_start=1
                     /transl_table=11
                     /product="CONSERVED HYPOTHETICAL PROTEIN"
                     /protein_id="CAD92866.1"
                     /db_xref="GI:31616766"
                     /translation="MTGSVDRPDQNRGERLMKSPGLDLVRRTLDEARAAARARGQDAG
                     RGRVASVASGRVAGRRRSWSGPGPDIRDPQPLGKAARELAKKRGWSVRVAEGMVLGQW
                     SAVVGHQIAEHARPTALNDGVLSVIAESTAWATQLRIMQAQLLAKIAAAVGNDVVRSL
                     KITGPAAPSWRKGPRHIAGRGPRDTYG"
     gene            5123. .7267
                     /gene="gyrB"
                     /locus_tag="Mb0005"
     CDS             5123. .7267
                     /gene="gyrB"
                     /locus_tag="Mb0005"
                     /EC_number="5.99.1.3"
                     /note="Mb0005, gyrB, len: 714 aa. Equivalent to Rv0005,
                     len: 714 aa, from Mycobacterium tuberculosis strain H37Rv,
                     (99.9% identity in 714 aa overlap). gyrB, DNA gyrase
                     subunit B (EC 5.99.1.3) (see citations below), equivalent,
                     except in N-terminus, to other Mycobacterial DNA GYRASES
                     SUBUNIT B e.g. T10005 from Mycobacterium leprae (697 aa);
                     Q9L7L3|GYRB_MYCPA from Mycobacterium avium subsp.
                     paratuberculosis (677 aa) (has its N-terminus shorter);
                     P48355|GYRB_MYCSM from Mycobacterium smegmatis (675 aa);
                     etc. Also highly similar to others e.g. T10969 from
                     Streptomyces coelicolor (686 aa); P50075|GYBS_STRSH from
```

Alignment Scores:

Pred. No.:	1.12e-39	Length:	343050
Score:	675.00	Matches:	132
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-684-215B-4 (1-132) x BX248334 (1-343050)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152008 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 152067

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152068 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 152127

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152128 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 152187

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152188 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 152247

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152248 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 152307

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     152308 GGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGT 152367

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db     152368 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 152403
```

```
RESULT 13
AR169152
LOCUS       AR169152                447 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 4 from patent US 6290969.
ACCESSION   AR169152
VERSION     AR169152.1  GI:17906927
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6290969-A 4 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Alignment Scores:

Pred. No.:	3.59e-42	Length:	447
Score:	670.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	99.26%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x AR169152 (1-447)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406
```

```
RESULT 14
AR182442
LOCUS       AR182442                447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6338852.
ACCESSION   AR182442
VERSION     AR182442.1  GI:20225649
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6338852-A 4 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                3.59e-42        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x AR182442 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 15
AR194825
LOCUS       AR194825                 447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6350456.
ACCESSION   AR194825
VERSION     AR194825.1  GI:20244262
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
  TITLE     Compositions and methods for the prevention and treatment of M.
            tuberculosis infection
  JOURNAL   Patent: US 6350456-A 4 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-42        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x AR194825 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 16
AR233097
LOCUS       AR233097                 447 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 4 from patent US 6458366.
ACCESSION   AR233097
VERSION     AR233097.1  GI:27275533
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6458366-A 4 01-OCT-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-42        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x AR233097 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
```

```
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 17
AR353302
LOCUS       AR353302                447 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 4 from patent US 6592877.
ACCESSION   AR353302
VERSION     AR353302.1  GI:33759108
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6592877-A 4 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                 3.59e-42        Length:          447
Score:                     670.00          Matches:         131
Percent Similarity:        99.24%          Conservative:    0
Best Local Similarity:     99.24%          Mismatches:      1
Query Match:               99.26%          Indels:          0
DB:                        6               Gaps:            0

US-09-684-215B-4 (1-132) x AR353302 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60

Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 18
AX429596
LOCUS       AX429596                447 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 4 from Patent EP1203817.
ACCESSION   AX429596
VERSION     AX429596.1  GI:21540845
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: EP 1203817-A 4 08-MAY-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                 3.59e-42        Length:          447
Score:                     670.00          Matches:         131
Percent Similarity:        99.24%          Conservative:    0
Best Local Similarity:     99.24%          Mismatches:      1
Query Match:               99.26%          Indels:          0
DB:                        6               Gaps:            0

US-09-684-215B-4 (1-132) x AX429596 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20

Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40

Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60

Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80

Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100

Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120

Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132

Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 19
AX832581
LOCUS       AX832581                447 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 4 from Patent EP1347055.
ACCESSION   AX832581
VERSION     AX832581.1  GI:39840631
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S., Twardzik,D.R. and Dillon,D.C.
```

```
TITLE      Compounds for immunotherapy and diagnosis of tuberculosis
JOURNAL    Patent: EP 1347055-A 4 24-SEP-2003;
           CORIXA CORPORATION (US)
FEATURES            Location/Qualifiers
     source         1..447
                    /organism="unidentified"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:               3.59e-42     Length:         447
Score:                   670.00       Matches:        131
Percent Similarity:      99.24%       Conservative:   0
Best Local Similarity:   99.24%       Mismatches:     1
Query Match:             99.26%       Indels:         0
DB:                      6            Gaps:           0

US-09-684-215B-4 (1-132) x AX832581 (1-447)

Qy        1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy       21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTCAT 130

Qy       41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy       61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy       81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy      101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
            |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db      311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy      121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
            ||||||||||||||||||||||||||||||||||||
Db      371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 20
BD006325
LOCUS       BD006325                447 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Compounds and methods for diagnosis of Tuberculosis.
ACCESSION   BD006325
VERSION     BD006325.1  GI:18634696
KEYWORDS    JP 2001500383-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
  TITLE     Compounds and methods for diagnosis of Tuberculosis
  JOURNAL   Patent: JP 2001500383-A 4 16-JAN-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001500383-A/4
            PD   16-JAN-2001
            PF   07-OCT-1997 JP 1998518432
            PR   11-OCT-1996 US    08/729622,13-MAR-1997 US    08/818111 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI   ANTONIO CAMPOS
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
```

```
            PC   C12N15/31,C07K14/35,C07K16/12,C12Q1/68,C12N15/62,G01N33/53 CC
             Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1..447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1..447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:               3.59e-42     Length:         447
Score:                   670.00       Matches:        131
Percent Similarity:      99.24%       Conservative:   0
Best Local Similarity:   99.24%       Mismatches:     1
Query Match:             99.26%       Indels:         0
DB:                      6            Gaps:           0

US-09-684-215B-4 (1-132) x BD006325 (1-447)

Qy        1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy       21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGTCACCCACCGTTCAT 130

Qy       41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy       61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy       81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy      101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
            |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db      311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy      121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
            ||||||||||||||||||||||||||||||||||||
Db      371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 21
BD006445
LOCUS       BD006445                447 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            Tuberculosis.
ACCESSION   BD006445
VERSION     BD006445.1  GI:18634816
KEYWORDS    JP 2001501832-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001501832-A 4 13-FEB-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001501832-A/4
            PD   13-FEB-2001
            PF   07-OCT-1997 JP 1998518456
            PR   11-OCT-1996 US    08/730510,13-MAR-1997 US    08/818112 PI
```

```
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI   ANTONIO CAMPOS
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC   C12N15/31,C07K14/35,A61K39/04,A61K48/00,A61K49/00,C12N15/62,
            PC   C07K19/00,
            PC   G01N33/50,G01N33/60,G01N33/569,C12N1/19,C12N1/20,C12N1/21, PC
             C12N5/10//
            PC   (C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-42        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x BD006445 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 22
BD069285
LOCUS       BD069285                 447 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            tuberculosis.
ACCESSION   BD069285
VERSION     BD069285.1  GI:22614888
KEYWORDS    JP 2001517069-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.H. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001517069-A 4 02-OCT-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001517069-A/4
            PD   02-OCT-2001
            PF   30-AUG-1996 JP 1997511464
            PR   01-SEP-1995 US    08/523436,22-SEP-1995 US    08/533634 PR
            22-MAR-1996 US    08/620874,05-JUN-1996 US    08/659683 PR
            12-JUL-1996 US    08/680574
            PI   STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
            PI   NETO,
            PI   RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
            C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
            C12N5/10,
            PC   C12N1/21//A61K39/04,(C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Compounds and methods for immunotherapy and diagnosis of  CC
                 tuberculosis
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-42        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x BD069285 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              |||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 23
BD205817
LOCUS       BD205817                 447 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compounds and methods for diagnosis of tuberculosis.
ACCESSION   BD205817
```

```
VERSION     BD205817.1  GI:33015587
KEYWORDS    JP 2002530050-A/4.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: JP 2002530050-A 4 17-SEP-2002;
            CORIXA CORP
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002530050-A/4
            PD   17-SEP-2002
            PF   17-FEB-1999 JP 2000532132
            PR   18-FEB-1998 US    09/024753,05-MAY-1998 US    09/072596 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS  PI
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK,MICHAEL J
            PI   LODES,
            PI   RONALD C HENDRICKSON
            PC   C12N15/09,C07H21/02,C07H21/04,C07K7/00,C07K14/35,C07K16/12, PC
              C07K17/00,
            PC   C07K19/00,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/04,C12Q1/
            PC   68,
            PC   G01N33/569,G01N33/68//A61K38/00,A61K39/04,A61K39/395,A61K39/
            PC   395,A61P31/06,
            PC   (C12N1/21,C12R1:19),C12N15/00,C12N5/00,A61K37/02 CC
            Compounds and methods for diagnosis of tuberculosis. FH   Key
                 Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Mycobacterium tuberculosis'.
FEATURES            Location/Qualifiers
     source         1. .447
                    /organism="Mycobacterium tuberculosis"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:1773"
ORIGIN

Alignment Scores:
Pred. No.:                3.59e-42        Length:         447
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x BD205817 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 130

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 190

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GTCCAACGCGTGGTCGGGAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 250

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 310

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db        311 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 370

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db        371 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 406

RESULT 24
AR169165
LOCUS       AR169165                1872 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 17 from patent US 6290969.
ACCESSION   AR169165
VERSION     AR169165.1  GI:17906947
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6290969-A 17 18-SEP-2001;
FEATURES            Location/Qualifiers
     source         1. .1872
                    /organism="unknown"
                    /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.49e-41        Length:         1872
Score:                    670.00          Matches:        131
Percent Similarity:       99.24%          Conservative:   0
Best Local Similarity:    99.24%          Mismatches:     1
Query Match:              99.26%          Indels:         0
DB:                       6               Gaps:           0

US-09-684-215B-4 (1-132) x AR169165 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 25
AR182455
LOCUS       AR182455                1872 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 17 from patent US 6338852.
ACCESSION   AR182455
VERSION     AR182455.1  GI:20225662
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6338852-A 17 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.49e-41      Length:          1872
Score:                    670.00        Matches:         131
Percent Similarity:       99.24%        Conservative:    0
Best Local Similarity:    99.24%        Mismatches:      1
Query Match:              99.26%        Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-4 (1-132) x AR182455 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 26
AR194838
LOCUS       AR194838               1872 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 17 from patent US 6350456.
ACCESSION   AR194838
VERSION     AR194838.1  GI:20244275
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
  TITLE     Compositions and methods for the prevention and treatment of M.
            tuberculosis infection
  JOURNAL   Patent: US 6350456-A 17 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.49e-41      Length:          1872
Score:                    670.00        Matches:         131
Percent Similarity:       99.24%        Conservative:    0
Best Local Similarity:    99.24%        Mismatches:      1
Query Match:              99.26%        Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-4 (1-132) x AR194838 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153

RESULT 27
AR233110
LOCUS       AR233110               1872 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 17 from patent US 6458366.
ACCESSION   AR233110
VERSION     AR233110.1  GI:27275546
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1872)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6458366-A 17 01-OCT-2002;
FEATURES             Location/Qualifiers
     source          1. .1872
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.49e-41      Length:          1872
Score:                    670.00        Matches:         131
Percent Similarity:       99.24%        Conservative:    0
Best Local Similarity:    99.24%        Mismatches:      1
Query Match:              99.26%        Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-4 (1-132) x AR233110 (1-1872)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817
```

```
Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db        1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||
Db        1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

RESULT 28
AR353315
LOCUS AR353315 1872 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 17 from patent US 6592877.
ACCESSION AR353315
VERSION AR353315.1 GI:33759121
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1872)
AUTHORS Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R., Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
TITLE Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL Patent: US 6592877-A 17 15-JUL-2003;
FEATURES Location/Qualifiers
source 1. .1872
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:

Pred. No.:	1.49e-41	Length:	1872
Score:	670.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	99.26%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x AR353315 (1-1872)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db        1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||
Db        1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

RESULT 29
AX429609
LOCUS AX429609 1872 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 17 from Patent EP1203817.
ACCESSION AX429609
VERSION AX429609.1 GI:21540858
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1
AUTHORS Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L., Vedvick,T.S. and Twardzik,D.R.
TITLE Compounds and methods for immunotherapy and diagnosis of tuberculosis
JOURNAL Patent: EP 1203817-A 17 08-MAY-2002;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .1872
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN

Alignment Scores:

Pred. No.:	1.49e-41	Length:	1872
Score:	670.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	99.26%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x AX429609 (1-1872)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy          41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy          61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy          81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy         101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
               |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db        1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy         121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
               ||||||||||||||||||||||||||||||||
Db        1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

RESULT 30
AX832594
LOCUS AX832594 1872 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 17 from Patent EP1347055.
ACCESSION AX832594
VERSION AX832594.1 GI:39840644
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1
AUTHORS Reed,S.G., Skeiky,Y.A., Campos-Neto,A., Houghton,R.L., Vedvick,T.S., Twardzik,D.R. and Dillon,D.C.
TITLE Compounds for immunotherapy and diagnosis of tuberculosis
JOURNAL Patent: EP 1347055-A 17 24-SEP-2003; CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .1872
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN

Alignment Scores:

Pred. No.:	1.49e-41	Length:	1872
Score:	670.00	Matches:	131
Percent Similarity:	99.24%	Conservative:	0
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	99.26%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-4 (1-132) x AX832594 (1-1872)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        758 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 817

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIleArgSerGlyGlyGlySerProThrValHis 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        818 GGGCAGGCGATGGCGATCGCGGGCCAAATCCGATCGGGTGGGGGGTCACCCACCGTTCAT 877

Qy         41 IleGlyProThrAlaPheLeuGlyLeuGlyValValAspAsnAsnGlyAsnGlyAlaArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 ATCGGGCCTACCGCCTTCCTCGGCTTGGGTGTTGTCGACAACAACGGCAACGGCGCACGA 937

Qy         61 ValGlnArgValValGlySerAlaProAlaAlaSerLeuGlyIleSerThrGlyAspVal 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        938 GTCCAACGCGTGGTCGGAAGCGCTCCGGCGGCAAGTCTCGGCATCTCCACCGGCGACGTG 997

Qy         81 IleThrAlaValAspGlyAlaProIleAsnSerAlaThrAlaMetAlaAspAlaLeuAsn 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        998 ATCACCGCGGTCGACGGCGCTCCGATCAACTCGGCCACCGCGATGGCGGACGCGCTTAAC 1057

Qy        101 GlyHisHisProGlyAspValIleSerValThrTrpGlnThrLysSerGlyGlyThrArg 120
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db       1058 GGGCATCATCCCGGTGACGTCATCTCGGTGAACTGGCAAACCAAGTCGGGCGGCACGCGT 1117

Qy        121 ThrGlyAsnValThrLeuAlaGluGlyProProAla 132
              ||||||||||||||||||||||||||||||||||||
Db       1118 ACAGGGAACGTGACATTGGCCGAGGGACCCCCGGCC 1153
```

Search completed: April 30, 2004, 02:24:00
Job time : 3854.74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:43:07 ; Search time 109.952 Seconds
(without alignments)
1234.209 Million cell updates/sec

Title: US-09-684-215B-17
Perfect score: 148
Sequence: 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13266/app_query.fasta_1.1180
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09684215_@CGN_1_1_370_runat_29042004_061306_13266
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	148	100.0	186	9	US-09-902-941-1879	Sequence 1879, Ap
2	148	100.0	186	9	US-09-849-626-1879	Sequence 1879, Ap
3	148	100.0	186	13	US-10-283-017-1879	Sequence 1879, Ap
4	148	100.0	186	15	US-10-017-754-1879	Sequence 1879, Ap
5	148	100.0	186	15	US-10-113-872-1879	Sequence 1879, Ap
6	148	100.0	399	9	US-09-757-417-46	Sequence 46, Appl
7	148	100.0	399	15	US-10-042-945-46	Sequence 46, Appl
8	148	100.0	447	15	US-10-193-002-4	Sequence 4, Appli
9	148	100.0	447	15	US-10-084-843-4	Sequence 4, Appli
10	148	100.0	447	15	US-10-098-732A-9	Sequence 9, Appli
11	148	100.0	675	9	US-09-759-143-822	Sequence 822, App
12	148	100.0	675	9	US-09-780-669-822	Sequence 822, App
13	148	100.0	675	9	US-09-822-827-822	Sequence 822, App
14	148	100.0	675	9	US-09-895-793-822	Sequence 822, App
15	148	100.0	675	9	US-09-895-814-822	Sequence 822, App
16	148	100.0	675	14	US-10-012-896-822	Sequence 822, App
17	148	100.0	675	15	US-10-144-678A-822	Sequence 822, App
18	148	100.0	675	15	US-10-294-025-822	Sequence 822, App
19	148	100.0	702	9	US-09-287-849-27	Sequence 27, Appl
20	148	100.0	702	15	US-10-359-460-27	Sequence 27, Appl
21	148	100.0	822	9	US-09-736-457-1862	Sequence 1862, Ap
22	148	100.0	822	9	US-09-902-941-1862	Sequence 1862, Ap
23	148	100.0	822	9	US-09-849-626-1862	Sequence 1862, Ap
24	148	100.0	822	13	US-10-283-017-1862	Sequence 1862, Ap
25	148	100.0	822	15	US-10-017-754-1862	Sequence 1862, Ap
26	148	100.0	822	15	US-10-113-872-1862	Sequence 1862, Ap
27	148	100.0	861	9	US-09-902-941-1877	Sequence 1877, Ap
28	148	100.0	861	9	US-09-849-626-1877	Sequence 1877, Ap
29	148	100.0	861	13	US-10-283-017-1877	Sequence 1877, Ap
30	148	100.0	861	15	US-10-017-754-1877	Sequence 1877, Ap
31	148	100.0	861	15	US-10-113-872-1877	Sequence 1877, Ap
32	148	100.0	894	13	US-09-878-722-236	Sequence 236, App
33	148	100.0	894	13	US-09-904-456-236	Sequence 236, App
34	148	100.0	900	9	US-09-735-705-353	Sequence 353, App
35	148	100.0	900	9	US-09-850-716A-353	Sequence 353, App
36	148	100.0	900	9	US-09-897-778-353	Sequence 353, App
37	148	100.0	900	13	US-10-007-700-353	Sequence 353, App
38	148	100.0	900	15	US-10-117-982-353	Sequence 353, App
39	148	100.0	900	16	US-10-313-986-353	Sequence 353, App
40	148	100.0	915	9	US-09-759-143-834	Sequence 834, App
41	148	100.0	915	9	US-09-780-669-834	Sequence 834, App
42	148	100.0	915	9	US-09-822-827-834	Sequence 834, App
43	148	100.0	915	9	US-09-895-793-834	Sequence 834, App
44	148	100.0	915	9	US-09-895-814-834	Sequence 834, App
45	148	100.0	915	14	US-10-012-896-834	Sequence 834, App

ALIGNMENTS

RESULT 1
US-09-902-941-1879
; Sequence 1879, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1879
; LENGTH: 186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1879

Alignment Scores:

Pred. No.:	1.99e-16	Length:	186
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-902-941-1879 (1-186)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 2
US-09-849-626-1879
; Sequence 1879, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1879
; LENGTH: 186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-1879

Alignment Scores:

Pred. No.:	1.99e-16	Length:	186
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-849-626-1879 (1-186)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 3
US-10-283-017-1879
; Sequence 1879, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C20
; CURRENT APPLICATION NUMBER: US/10/283,017
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1879
; LENGTH: 186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-017-1879

Alignment Scores:

Pred. No.:	1.99e-16	Length:	186
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-684-215B-17 (1-30) x US-10-283-017-1879 (1-186)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 4
US-10-017-754-1879
; Sequence 1879, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1879
; LENGTH: 186
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-1879

Alignment Scores:

```
Pred. No.:              1.99e-16        Length:         186
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-17 (1-30) x US-10-017-754-1879 (1-186)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 5
US-10-113-872-1879
; Sequence 1879, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
;  APPLICANT: Watanabe, Yoshihiro
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.478C19
;  CURRENT APPLICATION NUMBER: US/10/113,872
;  CURRENT FILING DATE:  2002-03-28
;  NUMBER OF SEQ ID NOS: 2011
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1879
;   LENGTH: 186
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-113-872-1879

Alignment Scores:
Pred. No.:              1.99e-16        Length:         186
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-17 (1-30) x US-10-113-872-1879 (1-186)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 6
US-09-757-417-46
; Sequence 46, Application US/09757417
; Patent No. US20020082216A1
; GENERAL INFORMATION:
;  APPLICANT: Fanger, Gary R.
;  APPLICANT:  Foy, Theresa M.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Reed, Steven G.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY, DIAGNOSIS AND MONITORING OF BREAST CANCER
;  FILE REFERENCE: 210121.479C1
;  CURRENT APPLICATION NUMBER: US/09/757,417
;  CURRENT FILING DATE:  2001-01-08
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
;   LENGTH: 399
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-757-417-46

Alignment Scores:
Pred. No.:              4.77e-16        Length:         399
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-17 (1-30) x US-09-757-417-46 (1-399)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 7
US-10-042-945-46
; Sequence 46, Application US/10042945
; Publication No. US20030045468A1
; GENERAL INFORMATION:
;  APPLICANT: Fling, Steven P.
;  APPLICANT:  Foy, Teresa M.
;  APPLICANT:  Clapper, Jonathan D.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Johnson, Jeffrey C.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY,
;  TITLE OF INVENTION:  DIAGNOSIS AND MONITORING OF BREAST CANCER
;  FILE REFERENCE: 210121.479C3
;  CURRENT APPLICATION NUMBER: US/10/042,945
;  CURRENT FILING DATE:  2002-01-08
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
;   LENGTH: 399
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-042-945-46

Alignment Scores:
Pred. No.:              4.77e-16        Length:         399
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-17 (1-30) x US-10-042-945-46 (1-399)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 8
US-10-193-002-4
; Sequence 4, Application US/10193002
```

```
; Publication No. US20030135026A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonia
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
;                            TUBERCULOSIS
;        NUMBER OF SEQUENCES: 350
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/193,002
;             FILING DATE: 10-Jul-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,596
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.417C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 447 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-193-002-4

Alignment Scores:
Pred. No.:                5.44e-16       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       15             Gaps:            0

US-09-684-215B-17 (1-30) x US-10-193-002-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 9
US-10-084-843-4
; Sequence 4, Application US/10084843
; Publication No. US20030143243A1
;   GENERAL INFORMATION:
;        APPLICANT: Reed, Steven G.
;                   Skeiky, Yasir A.W.
;                   Dillon, Davin C.
;                   Campos-Neto, Antonio
;                   Houghton, Raymond
;                   Vedvick, Thomas S.
;                   Twardzik, Daniel R.
;                   Lodes, Michael J.
;                   Hendrickson, Ronald C.
;        TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;                            AND DIAGNOSIS OF TUBERCULOSIS
;        NUMBER OF SEQUENCES: 355
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/084,843
;             FILING DATE: 25-Feb-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/072,967
;             FILING DATE: 05-MAY-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Maki, David J.
;             REGISTRATION NUMBER: 31,392
;             REFERENCE/DOCKET NUMBER: 210121.411C9
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 447 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-084-843-4

Alignment Scores:
Pred. No.:                5.44e-16       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       15             Gaps:            0

US-09-684-215B-17 (1-30) x US-10-084-843-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 10
US-10-098-732A-9
; Sequence 9, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir
;  APPLICANT:  Brannon, Mark
;  APPLICANT:  Guderian, Jeffrey
;  APPLICANT:  Corixa Corporation
```

; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTBRa12 (MTB32A
; OTHER INFORMATION: C-terminus)
US-10-098-732A-9

Alignment Scores:

Pred. No.:	5.44e-16	Length:	447
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-17 (1-30) x US-10-098-732A-9 (1-447)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100
```

RESULT 11
US-09-759-143-822
; Sequence 822, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-822

Alignment Scores:

Pred. No.:	8.73e-16	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-759-143-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 12
US-09-780-669-822
; Sequence 822, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-822

Alignment Scores:

Pred. No.:	8.73e-16	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-780-669-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 13

US-09-822-827-822
; Sequence 822, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-822

Alignment Scores:

Pred. No.:	8.73e-16	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-822-827-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 14
US-09-895-793-822
; Sequence 822, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-822

Alignment Scores:

Pred. No.:	8.73e-16	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-895-793-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 15
US-09-895-814-822
; Sequence 822, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-822

Alignment Scores:

Pred. No.:	8.73e-16	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-895-814-822 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 16
US-10-012-896-822
; Sequence 822, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals de Bassols, Carlota
;  APPLICANT:  Foy, Teresa
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Wantanabe, Yoshihiro
;  APPLICANT:  Meagher, Madeleine Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C27
;  CURRENT APPLICATION NUMBER: US/10/012,896
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 1011
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-012-896-822

Alignment Scores:
Pred. No.:                8.73e-16        Length:          675
Score:                    148.00          Matches:         30
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       14              Gaps:            0

US-09-684-215B-17 (1-30) x US-10-012-896-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 17
US-10-144-678A-822
; Sequence 822, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A. W.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Hural, John
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Vinals y de Bassols, Carlota
;  APPLICANT:  Foy, Teresa M.
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Deng, Ta
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C28
;  CURRENT APPLICATION NUMBER: US/10/144,678A
;  CURRENT FILING DATE:  2002-08-12
;  NUMBER OF SEQ ID NOS: 1033
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-144-678A-822

Alignment Scores:
Pred. No.:                8.73e-16        Length:          675
Score:                    148.00          Matches:         30
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       15              Gaps:            0

US-09-684-215B-17 (1-30) x US-10-144-678A-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 18
US-10-294-025-822
; Sequence 822, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Kalos, Michael D.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C29
;  CURRENT APPLICATION NUMBER: US/10/294,025
;  CURRENT FILING DATE:  2002-11-12
;  NUMBER OF SEQ ID NOS: 1038
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;   LENGTH: 675
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-294-025-822
```

```
Alignment Scores:
Pred. No.:              8.73e-16        Length:         675
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0

US-09-684-215B-17 (1-30) x US-10-294-025-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              |||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 19
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. US20020009459A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION:  Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE:  014058-009020US
;  CURRENT APPLICATION NUMBER: US/09/287,849
;  CURRENT FILING DATE:  1999-04-07
;  PRIOR APPLICATION NUMBER:  US 08/818,112
;  PRIOR FILING DATE:  1997-03-13
;  PRIOR APPLICATION NUMBER:  US 08/942,578
;  PRIOR FILING DATE:  1997-10-01
;  PRIOR APPLICATION NUMBER:  US 09/025,197
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 09/056,556
;  PRIOR FILING DATE:  1998-04-07
;  PRIOR APPLICATION NUMBER:  US 09/223,040
;  PRIOR FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  46
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 702
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
;   OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
;   OTHER INFORMATION: frame 1
;   NAME/KEY: CDS
;   LOCATION: (1)..(693)
;   OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
;   OTHER INFORMATION: reading frame 1
;   NAME/KEY: CDS
;   LOCATION: (2)..(700)
;   OTHER INFORMATION: reading frame 2
;   NAME/KEY: CDS
;   LOCATION: (3)..(701)
;   OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:
Pred. No.:              9.13e-16        Length:         702
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-684-215B-17 (1-30) x US-09-287-849-27 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              |||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114

RESULT 20
US-10-359-460-27
; Sequence 27, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION: Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE: 014058-009020US
;  CURRENT APPLICATION NUMBER: US/10/359,460
;  CURRENT FILING DATE:  2003-02-05
;  PRIOR APPLICATION NUMBER: US/09/287,849
;  PRIOR FILING DATE: 1999-04-07
;  PRIOR APPLICATION NUMBER: US 08/818,112
;  PRIOR FILING DATE: 1997-03-13
;  PRIOR APPLICATION NUMBER: US 08/942,578
;  PRIOR FILING DATE: 1997-10-01
;  PRIOR APPLICATION NUMBER: US 09/025,197
;  PRIOR FILING DATE: 1998-02-18
;  PRIOR APPLICATION NUMBER: US 09/056,556
;  PRIOR FILING DATE: 1998-04-07
;  PRIOR APPLICATION NUMBER: US 09/223,040
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 702
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
;   OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
;   OTHER INFORMATION: frame 1
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(693)
;   OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
;   OTHER INFORMATION: reading frame 1
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(700)
;   OTHER INFORMATION: reading frame 2
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(701)
;   OTHER INFORMATION: reading frame 3
US-10-359-460-27

Alignment Scores:
Pred. No.:              9.13e-16        Length:         702
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     15              Gaps:           0
```

US-09-684-215B-17 (1-30) x US-10-359-460-27 (1-702)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114
```

RESULT 21
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1862

Alignment Scores:

Pred. No.:	1.09e-15	Length:	822
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-736-457-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 22
US-09-902-941-1862
; Sequence 1862, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1862

Alignment Scores:

Pred. No.:	1.09e-15	Length:	822
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-902-941-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 23
US-09-849-626-1862
; Sequence 1862, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-1862

Alignment Scores:

Pred. No.:	1.09e-15	Length:	822
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-849-626-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 24
US-10-283-017-1862
; Sequence 1862, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C20
; CURRENT APPLICATION NUMBER: US/10/283,017
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-017-1862

Alignment Scores:

Pred. No.:	1.09e-15	Length:	822
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-684-215B-17 (1-30) x US-10-283-017-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 25
US-10-017-754-1862
; Sequence 1862, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-1862

Alignment Scores:

Pred. No.:	1.09e-15	Length:	822
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-17 (1-30) x US-10-017-754-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 26
US-10-113-872-1862
; Sequence 1862, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-1862

Alignment Scores:

Pred. No.:	1.09e-15	Length:	822
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-17 (1-30) x US-10-113-872-1862 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 27
US-09-902-941-1877
; Sequence 1877, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1877
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-1877

Alignment Scores:

Pred. No.:	1.15e-15	Length:	861
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-902-941-1877 (1-861)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 28
US-09-849-626-1877
; Sequence 1877, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1877
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-1877

Alignment Scores:

Pred. No.:	1.15e-15	Length:	861
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-849-626-1877 (1-861)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 29
US-10-283-017-1877
; Sequence 1877, Application US/10283017
; Publication No. US20030211510A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C20
; CURRENT APPLICATION NUMBER: US/10/283,017
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1877
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-017-1877

Alignment Scores:

Pred. No.:	1.15e-15	Length:	861
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-684-215B-17 (1-30) x US-10-283-017-1877 (1-861)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 30
US-10-017-754-1877
; Sequence 1877, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1877
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-1877

Alignment Scores:

Pred. No.:	1.15e-15	Length:	861
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0

US-09-684-215B-17 (1-30) x US-10-017-754-1877 (1-861)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

Search completed: April 30, 2004, 05:10:22
Job time : 111.202 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:28:57 ; Search time 15.3589 Seconds
(without alignments)
1083.969 Million cell updates/sec

Title: US-09-684-215B-17
Perfect score: 148
Sequence: 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061306_13249/app_query.fasta_1.1180
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_115_@runat_29042004_061306_13249 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	148	100.0	447	3	US-08-818-112-4	Sequence 4, Appli
2	148	100.0	447	4	US-08-818-111-4	Sequence 4, Appli
3	148	100.0	447	4	US-09-056-556-4	Sequence 4, Appli
4	148	100.0	447	4	US-09-072-596-4	Sequence 4, Appli
5	148	100.0	447	4	US-09-072-967-4	Sequence 4, Appli
6	148	100.0	675	4	US-09-636-215-822	Sequence 822, App
7	148	100.0	675	4	US-09-685-166A-822	Sequence 822, App
8	148	100.0	702	4	US-09-287-849-27	Sequence 27, Appl
9	148	100.0	822	4	US-09-736-457-1862	Sequence 1862, Ap
10	148	100.0	900	4	US-09-643-597-353	Sequence 353, App
11	148	100.0	900	4	US-09-606-421B-353	Sequence 353, App
12	148	100.0	915	4	US-09-636-215-834	Sequence 834, App
13	148	100.0	915	4	US-09-685-166A-834	Sequence 834, App
14	148	100.0	945	4	US-09-736-457-1861	Sequence 1861, Ap
15	148	100.0	1012	4	US-09-643-597-351	Sequence 351, App
16	148	100.0	1012	4	US-09-606-421B-351	Sequence 351, App
17	148	100.0	1203	4	US-09-636-215-851	Sequence 851, App
18	148	100.0	1203	4	US-09-685-166A-851	Sequence 851, App
19	148	100.0	1464	4	US-09-620-412C-348	Sequence 348, App
20	148	100.0	1464	4	US-09-598-419-348	Sequence 348, App
21	148	100.0	1557	4	US-09-620-412C-332	Sequence 332, App
22	148	100.0	1557	4	US-09-598-419-332	Sequence 332, App
23	148	100.0	1578	4	US-09-556-877-188	Sequence 188, App
24	148	100.0	1578	4	US-09-620-412C-188	Sequence 188, App
25	148	100.0	1578	4	US-09-598-419-188	Sequence 188, App
26	148	100.0	1752	4	US-09-620-412C-352	Sequence 352, App
27	148	100.0	1752	4	US-09-598-419-352	Sequence 352, App
28	148	100.0	1758	4	US-09-620-412C-336	Sequence 336, App
29	148	100.0	1758	4	US-09-598-419-336	Sequence 336, App
30	148	100.0	1860	4	US-09-620-412C-308	Sequence 308, App
31	148	100.0	1860	4	US-09-598-419-308	Sequence 308, App
32	148	100.0	1872	3	US-08-818-112-17	Sequence 17, Appl
33	148	100.0	1872	4	US-08-818-111-17	Sequence 17, Appl
34	148	100.0	1872	4	US-09-056-556-17	Sequence 17, Appl
35	148	100.0	1872	4	US-09-072-596-17	Sequence 17, Appl
36	148	100.0	1872	4	US-09-072-967-17	Sequence 17, Appl
37	148	100.0	1896	4	US-09-620-412C-324	Sequence 324, App
38	148	100.0	1896	4	US-09-598-419-324	Sequence 324, App
39	148	100.0	1941	4	US-09-620-412C-316	Sequence 316, App
40	148	100.0	1941	4	US-09-598-419-316	Sequence 316, App
41	148	100.0	1965	4	US-09-620-412C-340	Sequence 340, App
42	148	100.0	1965	4	US-09-598-419-340	Sequence 340, App
43	148	100.0	2052	4	US-09-620-412C-356	Sequence 356, App
44	148	100.0	2052	4	US-09-598-419-356	Sequence 356, App
45	148	100.0	2076	4	US-09-620-412C-312	Sequence 312, App

ALIGNMENTS

RESULT 1
US-08-818-112-4
; Sequence 4, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELECOMMUNICATION INFORMATION:

```
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  447 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-818-112-4

Alignment Scores:
Pred. No.:                5.52e-16        Length:         447
Score:                    148.00          Matches:        30
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       3               Gaps:           0

US-09-684-215B-17 (1-30) x US-08-818-112-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 2
US-08-818-111-4
; Sequence 4, Application US/08818111
; Patent No. 6338852
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/818,111
;      FILING DATE:  13-MAR-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-818-111-4

Alignment Scores:
Pred. No.:                5.52e-16        Length:         447
Score:                    148.00          Matches:        30
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-17 (1-30) x US-08-818-111-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 3
US-09-056-556-4
; Sequence 4, Application US/09056556
; Patent No. 6350456
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND              TREAT
;    NUMBER OF SEQUENCES:  241
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/056,556
;      FILING DATE:  07-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.457
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-056-556-4

Alignment Scores:
Pred. No.:                5.52e-16        Length:         447
Score:                    148.00          Matches:        30
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-17 (1-30) x US-09-056-556-4 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70
```

```
Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 4
US-09-072-596-4
; Sequence 4, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-4

Alignment Scores:
Pred. No.:              5.52e-16        Length:         447
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-17 (1-30) x US-09-072-596-4 (1-447)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 5
US-09-072-967-4
; Sequence 4, Application US/09072967
; Patent No. 6592877
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonio
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
;    TITLE OF INVENTION:  AND DIAGNOSIS OF TUBERCULOSIS
;    NUMBER OF SEQUENCES:  355
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,967
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.411C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  447 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-967-4

Alignment Scores:
Pred. No.:              5.52e-16        Length:         447
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-17 (1-30) x US-09-072-967-4 (1-447)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 6
US-09-636-215-822
; Sequence 822, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
```

```
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;    LENGTH: 675
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-636-215-822

Alignment Scores:
Pred. No.:                9.28e-16        Length:         675
Score:                    148.00          Matches:        30
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-17 (1-30) x US-09-636-215-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 7
US-09-685-166A-822
; Sequence 822, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 822
;    LENGTH: 675
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-685-166A-822

Alignment Scores:
Pred. No.:                9.28e-16        Length:         675
Score:                    148.00          Matches:        30
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       4               Gaps:           0

US-09-684-215B-17 (1-30) x US-09-685-166A-822 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 8
US-09-287-849-27
; Sequence 27, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Alderson, Mark
;  APPLICANT:  Campos-Neto, Antonio
;  APPLICANT:  Corixa Corporation
;  TITLE OF INVENTION:  Fusion Protiens of Mycobacterium tuberculosis Antigens
;  TITLE OF INVENTION:  and Their Uses
;  FILE REFERENCE:  014058-009020US
;  CURRENT APPLICATION NUMBER: US/09/287,849
;  CURRENT FILING DATE:  1999-04-07
;  PRIOR APPLICATION NUMBER:  US 08/818,112
;  PRIOR FILING DATE:  1997-03-13
;  PRIOR APPLICATION NUMBER:  US 08/942,578
;  PRIOR FILING DATE:  1997-10-01
;  PRIOR APPLICATION NUMBER:  US 09/025,197
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 09/056,556
;  PRIOR FILING DATE:  1998-04-07
;  PRIOR APPLICATION NUMBER:  US 09/223,040
;  PRIOR FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  46
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 27
;    LENGTH: 702
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
;    OTHER INFORMATION: protein Ra12-DPPD (designated Mtb24), reading
;    OTHER INFORMATION: frame 1
;    NAME/KEY: CDS
;    LOCATION: (1)..(693)
;    OTHER INFORMATION: bi-fusion protein Ra12-DPPD (designated Mtb24),
;    OTHER INFORMATION: reading frame 1
;    NAME/KEY: CDS
;    LOCATION: (2)..(700)
;    OTHER INFORMATION: reading frame 2
;    NAME/KEY: CDS
;    LOCATION: (3)..(701)
;    OTHER INFORMATION: reading frame 3
US-09-287-849-27

Alignment Scores:
Pred. No.:                9.75e-16        Length:         702
Score:                    148.00          Matches:        30
```

```
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     4              Gaps:          0

US-09-684-215B-17 (1-30) x US-09-287-849-27 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114

RESULT 9
US-09-736-457-1862
; Sequence 1862, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Lodes, Michael A.
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Vedvick, Tom
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Retter, Marc
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Wang, Aijun
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.478C15
;  CURRENT APPLICATION NUMBER: US/09/736,457
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  1864
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 1862
;   LENGTH: 822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1862

Alignment Scores:
Pred. No.:              1.19e-15       Length:        822
Score:                  148.00         Matches:       30
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     4              Gaps:          0

US-09-684-215B-17 (1-30) x US-09-736-457-1862 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 10
US-09-643-597-353
; Sequence 353, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-353

Alignment Scores:
Pred. No.:              1.33e-15       Length:        900
Score:                  148.00         Matches:       30
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     4              Gaps:          0

US-09-684-215B-17 (1-30) x US-09-643-597-353 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 11
US-09-606-421B-353
; Sequence 353, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 353
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-606-421B-353

Alignment Scores:
Pred. No.:              1.33e-15       Length:        900
Score:                  148.00         Matches:       30
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     4              Gaps:          0

US-09-684-215B-17 (1-30) x US-09-606-421B-353 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
                ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 12
US-09-636-215-834
; Sequence 834, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-636-215-834

Alignment Scores:

Pred. No.:	1.36e-15	Length:	915
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-636-215-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
                ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 13
US-09-685-166A-834
; Sequence 834, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 834
;   LENGTH: 915
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-834

Alignment Scores:

Pred. No.:	1.36e-15	Length:	915
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-685-166A-834 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
                ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 14
US-09-736-457-1861
; Sequence 1861, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1861
;   LENGTH: 945
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1861

Alignment Scores:

Pred. No.:	1.42e-15	Length:	945
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

```
Query Match:            100.00%       Indels:         0
DB:                     4             Gaps:           0

US-09-684-215B-17 (1-30) x US-09-736-457-1861 (1-945)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 15
US-09-643-597-351
; Sequence 351, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 351
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-351

Alignment Scores:
Pred. No.:              1.55e-15      Length:         1012
Score:                  148.00        Matches:        30
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     4             Gaps:           0

US-09-684-215B-17 (1-30) x US-09-643-597-351 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATC 156

RESULT 16
US-09-606-421B-351
; Sequence 351, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE:  2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 351
;   LENGTH: 1012
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-606-421B-351

Alignment Scores:
Pred. No.:              1.55e-15      Length:         1012
Score:                  148.00        Matches:        30
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     4             Gaps:           0

US-09-684-215B-17 (1-30) x US-09-606-421B-351 (1-1012)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 126

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db        127 GGGCAGGCGATGGCGATCGCGGGCCAGATC 156

RESULT 17
US-09-636-215-851
; Sequence 851, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-636-215-851

Alignment Scores:
Pred. No.:              1.92e-15      Length:         1203
Score:                  148.00        Matches:        30
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
```

```
DB:                    4              Gaps:          0

US-09-684-215B-17 (1-30) x US-09-636-215-851 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 18
US-09-685-166A-851
; Sequence 851, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 851
;   LENGTH: 1203
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-685-166A-851

Alignment Scores:
Pred. No.:              1.92e-15       Length:          1203
Score:                  148.00         Matches:         30
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     4              Gaps:            0

US-09-684-215B-17 (1-30) x US-09-685-166A-851 (1-1203)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 19
US-09-620-412C-348
; Sequence 348, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
;   LENGTH: 1464
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-348

Alignment Scores:
Pred. No.:              2.47e-15       Length:          1464
Score:                  148.00         Matches:         30
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     4              Gaps:            0

US-09-684-215B-17 (1-30) x US-09-620-412C-348 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 20
US-09-598-419-348
; Sequence 348, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 348
;   LENGTH: 1464
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-348

Alignment Scores:
Pred. No.:              2.47e-15       Length:          1464
Score:                  148.00         Matches:         30
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     4              Gaps:            0

US-09-684-215B-17 (1-30) x US-09-598-419-348 (1-1464)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 21
US-09-620-412C-332
; Sequence 332, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 332
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Chlymadia trachomatis
US-09-620-412C-332

Alignment Scores:

Pred. No.:	2.67e-15	Length:	1557
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-620-412C-332 (1-1557)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 22
US-09-598-419-332
; Sequence 332, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 332
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Chlymadia trachomatis
US-09-598-419-332

Alignment Scores:

Pred. No.:	2.67e-15	Length:	1557
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-598-419-332 (1-1557)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 23
US-09-556-877-188
; Sequence 188, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 188
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-556-877-188

Alignment Scores:

Pred. No.:	2.71e-15	Length:	1578
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-556-877-188 (1-1578)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 24
US-09-620-412C-188
; Sequence 188, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 188
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-620-412C-188

Alignment Scores:

Pred. No.:	2.71e-15	Length:	1578
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-620-412C-188 (1-1578)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

```
RESULT 25
US-09-598-419-188
; Sequence 188, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 188
;   LENGTH: 1578
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-598-419-188

Alignment Scores:
Pred. No.:              2.71e-15      Length:          1578
Score:                  148.00        Matches:         30
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     4             Gaps:            0

US-09-684-215B-17 (1-30) x US-09-598-419-188 (1-1578)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 26
US-09-620-412C-352
; Sequence 352, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 352
;   LENGTH: 1752
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-352

Alignment Scores:
Pred. No.:              3.09e-15      Length:          1752
Score:                  148.00        Matches:         30
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     4             Gaps:            0

US-09-684-215B-17 (1-30) x US-09-620-412C-352 (1-1752)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 27
US-09-598-419-352
; Sequence 352, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Scholler, John
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C6
;  CURRENT APPLICATION NUMBER: US/09/598,419
;  CURRENT FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 357
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 352
;   LENGTH: 1752
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-598-419-352

Alignment Scores:
Pred. No.:              3.09e-15      Length:          1752
Score:                  148.00        Matches:         30
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     4             Gaps:            0

US-09-684-215B-17 (1-30) x US-09-598-419-352 (1-1752)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 28
US-09-620-412C-336
; Sequence 336, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 336
;   LENGTH: 1758
;   TYPE: DNA
;   ORGANISM: Chlymadia trachomatis
US-09-620-412C-336

Alignment Scores:
Pred. No.:              3.11e-15      Length:          1758
Score:                  148.00        Matches:         30
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     4             Gaps:            0

US-09-684-215B-17 (1-30) x US-09-620-412C-336 (1-1758)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81
```

```
Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 29
US-09-598-419-336
; Sequence 336, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 336
; LENGTH: 1758
; TYPE: DNA
; ORGANISM: Chlymadia trachomatis
US-09-598-419-336

Alignment Scores:

Pred. No.:	3.11e-15	Length:	1758
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-598-419-336 (1-1758)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 30
US-09-620-412C-308
; Sequence 308, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 308
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-308

Alignment Scores:

Pred. No.:	3.34e-15	Length:	1860
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x US-09-620-412C-308 (1-1860)

```
Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

Search completed: April 30, 2004, 05:01:08
Job time : 17.3589 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:32:52 ; Search time 860.167 Seconds
(without alignments)
1511.672 Million cell updates/sec

Title: US-09-684-215B-17
Perfect score: 148
Sequence: 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13185/app_query.fasta_1.1180
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_6034_@runat_29042004_061304_13185 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	148	100.0	186	6	AX369169	AX369169 Sequence
2	148	100.0	447	6	AR169152	AR169152 Sequence
3	148	100.0	447	6	AR182442	AR182442 Sequence
4	148	100.0	447	6	AR194825	AR194825 Sequence
5	148	100.0	447	6	AR233097	AR233097 Sequence
6	148	100.0	447	6	AR353302	AR353302 Sequence
7	148	100.0	447	6	AX429596	AX429596 Sequence
8	148	100.0	447	6	AX832581	AX832581 Sequence
9	148	100.0	447	6	BD006325	BD006325 Compounds
10	148	100.0	447	6	BD006445	BD006445 Compounds
11	148	100.0	447	6	BD069285	BD069285 Compounds
12	148	100.0	447	6	BD205817	BD205817 Compounds
13	148	100.0	675	6	AR261272	AR261272 Sequence
14	148	100.0	675	6	AR400535	AR400535 Sequence
15	148	100.0	675	6	AR405802	AR405802 Sequence
16	148	100.0	675	6	AX201049	AX201049 Sequence
17	148	100.0	675	6	AX267848	AX267848 Sequence
18	148	100.0	702	6	BD251334	BD251334 Fused pro
19	148	100.0	702	6	AR403747	AR403747 Sequence
20	148	100.0	822	6	AR277645	AR277645 Sequence
21	148	100.0	822	6	AX369152	AX369152 Sequence
22	148	100.0	861	6	AX369167	AX369167 Sequence
23	148	100.0	894	6	AX351489	AX351489 Sequence
24	148	100.0	900	6	AR220690	AR220690 Sequence
25	148	100.0	900	6	AX365960	AX365960 Sequence
26	148	100.0	915	6	AR261281	AR261281 Sequence
27	148	100.0	915	6	AR400544	AR400544 Sequence
28	148	100.0	915	6	AR405811	AR405811 Sequence
29	148	100.0	915	6	AX201061	AX201061 Sequence
30	148	100.0	915	6	AX267860	AX267860 Sequence
31	148	100.0	945	6	AR277644	AR277644 Sequence
32	148	100.0	945	6	AX369151	AX369151 Sequence
33	148	100.0	1012	6	AR220689	AR220689 Sequence
34	148	100.0	1012	6	AX365958	AX365958 Sequence
35	148	100.0	1068	6	BD274032	BD274032 Sequences
36	148	100.0	1068	6	AX005788	AX005788 Sequence
37	148	100.0	1143	6	BD274033	BD274033 Sequences
38	148	100.0	1143	6	AX005790	AX005790 Sequence
39	148	100.0	1155	6	AX369165	AX369165 Sequence
40	148	100.0	1203	6	AR400556	AR400556 Sequence
41	148	100.0	1203	6	AR405823	AR405823 Sequence
42	148	100.0	1203	6	AX201078	AX201078 Sequence
43	148	100.0	1203	6	AX267877	AX267877 Sequence
44	148	100.0	1464	6	AR229410	AR229410 Sequence
45	148	100.0	1464	6	AR322028	AR322028 Sequence

ALIGNMENTS

RESULT 1

```
AX369169
LOCUS       AX369169                186 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 1879 from Patent WO0204514.
ACCESSION   AX369169
VERSION     AX369169.1  GI:18857178
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
            Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
            Mcnabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0204514-A 1879 17-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .186
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.6e-14       Length:          186
Score:                    148.00        Matches:         30
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-17 (1-30) x AX369169 (1-186)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 2
AR169152
LOCUS       AR169152                447 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 4 from patent US 6290969.
ACCESSION   AR169152
VERSION     AR169152.1  GI:17906927
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6290969-A 4 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.18e-13      Length:          447
Score:                    148.00        Matches:         30
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-17 (1-30) x AR169152 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 3
AR182442
LOCUS       AR182442                447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6338852.
ACCESSION   AR182442
VERSION     AR182442.1  GI:20225649
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6338852-A 4 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.18e-13      Length:          447
Score:                    148.00        Matches:         30
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-17 (1-30) x AR182442 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 4
AR194825
LOCUS       AR194825                447 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6350456.
ACCESSION   AR194825
VERSION     AR194825.1  GI:20244262
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W. and Dillon,D.C.
  TITLE     Compositions and methods for the prevention and treatment of M.
            tuberculosis infection
  JOURNAL   Patent: US 6350456-A 4 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:                1.18e-13      Length:          447
Score:                    148.00        Matches:         30
```

```
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-17 (1-30) x AR194825 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 5
AR233097
LOCUS       AR233097                447 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 4 from patent US 6458366.
ACCESSION   AR233097
VERSION     AR233097.1  GI:27275533
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: US 6458366-A 4 01-OCT-2002;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.18e-13       Length:        447
Score:                  148.00         Matches:       30
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-17 (1-30) x AR233097 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 6
AR353302
LOCUS       AR353302                447 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 4 from patent US 6592877.
ACCESSION   AR353302
VERSION     AR353302.1  GI:33759108
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Campos-Neto,A., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: US 6592877-A 4 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.18e-13       Length:        447
Score:                  148.00         Matches:       30
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-17 (1-30) x AR353302 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 7
AX429596
LOCUS       AX429596                447 bp    DNA     linear   PAT 21-JUN-2002
DEFINITION  Sequence 4 from Patent EP1203817.
ACCESSION   AX429596
VERSION     AX429596.1  GI:21540845
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
            tuberculosis
  JOURNAL   Patent: EP 1203817-A 4 08-MAY-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:              1.18e-13       Length:        447
Score:                  148.00         Matches:       30
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-684-215B-17 (1-30) x AX429596 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 8
AX832581
LOCUS       AX832581                447 bp    DNA     linear   PAT 12-DEC-2003
DEFINITION  Sequence 4 from Patent EP1347055.
ACCESSION   AX832581
VERSION     AX832581.1  GI:39840631
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
```

```
REFERENCE   1
  AUTHORS   Reed,S.G., Skeiky,Y.A., Campos-Neto,A., Houghton,R.L.,
            Vedvick,T.S., Twardzik,D.R. and Dillon,D.C.
  TITLE     Compounds for immunotherapy and diagnosis of tuberculosis
  JOURNAL   Patent: EP 1347055-A 4 24-SEP-2003;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                1.18e-13       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-17 (1-30) x AX832581 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 9
BD006325
LOCUS       BD006325                447 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Compounds and methods for diagnosis of Tuberculosis.
ACCESSION   BD006325
VERSION     BD006325.1  GI:18634696
KEYWORDS    JP 2001500383-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
  TITLE     Compounds and methods for diagnosis of Tuberculosis
  JOURNAL   Patent: JP 2001500383-A 4 16-JAN-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001500383-A/4
            PD   16-JAN-2001
            PF   07-OCT-1997 JP 1998518432
            PR   11-OCT-1996 US    08/729622,13-MAR-1997 US    08/818111 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI   ANTONIO CAMPOS
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC   C12N15/31,C07K14/35,C07K16/12,C1?Q1/68,C12N15/62,G01N33/53 CC
             Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                1.18e-13       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-17 (1-30) x BD006325 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 10
BD006445
LOCUS       BD006445                447 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            Tuberculosis.
ACCESSION   BD006445
VERSION     BD006445.1  GI:18634816
KEYWORDS    JP 2001501832-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R. and Lodes,M.J.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001501832-A 4 13-FEB-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001501832-A/4
            PD   13-FEB-2001
            PF   07-OCT-1997 JP 1998518456
            PR   11-OCT-1996 US    08/730510,13-MAR-1997 US    08/818112 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON, PI   ANTONIO CAMPOS
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK, PI
            MICHAEL J LODES
            PC   C12N15/31,C07K14/35,A61K39/04,A61K48/00,A61K49/00,C12N15/62,
            PC   C07K19/00,
            PC   G01N33/50,G01N33/60,G01N33/569,C12N1/19,C12N1/20,C12N1/21, PC
             C12N5/10//
            PC   (C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                1.18e-13       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       6              Gaps:            0

US-09-684-215B-17 (1-30) x BD006445 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
```

```
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 11
BD069285
LOCUS       BD069285                447 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Compounds and methods for immunotherapy and diagnosis of
            tuberculosis.
ACCESSION   BD069285
VERSION     BD069285.1  GI:22614888
KEYWORDS    JP 2001517069-A/4.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.H. and Twardzik,D.R.
  TITLE     Compounds and methods for immunotherapy and diagnosis of
  JOURNAL   Patent: JP 2001517069-A 4 02-OCT-2001;
            CORIXA CORP
COMMENT     OS   Unidentified
            PN   JP 2001517069-A/4
            PD   02-OCT-2001
            PF   30-AUG-1996 JP 1997511464
            PR   01-SEP-1995 US    08/523436,22-SEP-1995 US    08/533634 PR
            22-MAR-1996 US    08/620874,05-JUN-1996 US    08/659683 PR
            12-JUL-1996 US    08/680574
            PI   STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS
            PI   NETO,
            PI   RAYMOND HOUGHTON,THOMAS H VEDVICK,DANIEL R TWARDZIK PC
            C12N15/31,C07K14/35,A61K38/16,C12N15/62,G01N33/569,C12Q1/68, PC
            C12N5/10,
            PC   C12N1/21//A61K39/04,(C12N1/21,C12R1:19)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Compounds and methods for immunotherapy and diagnosis of  CC
                 tuberculosis
            FH   Key             Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

Alignment Scores:
Pred. No.:                 1.18e-13        Length:          447
Score:                     148.00          Matches:         30
Percent Similarity:        100.00%         Conservative:    0
Best Local Similarity:     100.00%         Mismatches:      0
Query Match:               100.00%         Indels:          0
DB:                        6               Gaps:            0

US-09-684-215B-17 (1-30) x BD069285 (1-447)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 12
BD205817
LOCUS       BD205817                447 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Compounds and methods for diagnosis of tuberculosis.
ACCESSION   BD205817
VERSION     BD205817.1  GI:33015587
KEYWORDS    JP 2002530050-A/4.
SOURCE      Mycobacterium tuberculosis
  ORGANISM  Mycobacterium tuberculosis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
            tuberculosis complex.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Reed,S.G., Skeiky,Y.A.W., Dillon,D.C., Neto,A.C., Houghton,R.,
            Vedvick,T.S., Twardzik,D.R., Lodes,M.J. and Hendrickson,R.C.
  TITLE     Compounds and methods for diagnosis of tuberculosis
  JOURNAL   Patent: JP 2002530050-A 4 17-SEP-2002;
            CORIXA CORP
COMMENT     OS   Mycobacterium tuberculosis
            PN   JP 2002530050-A/4
            PD   17-SEP-2002
            PF   17-FEB-1999 JP 2000532132
            PR   18-FEB-1998 US    09/024753,05-MAY-1998 US    09/072596 PI
            STEVEN G REED,YASIR A W SKEIKY,DAVIN C DILLON,ANTONIO CAMPOS  PI
            NETO,
            PI   RAYMOND HOUGHTON,THOMAS S VEDVICK,DANIEL R TWARDZIK,MICHAEL J
            PI   LODES,
            PI   RONALD C HENDRICKSON
            PC   C12N15/09,C07H21/02,C07H21/04,C07K7/00,C07K14/35,C07K16/12, PC
              C07K17/00,
            PC   C07K19/00,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/04,C12Q1/
            PC   68,
            PC   G01N33/569,G01N33/68//A61K38/00,A61K39/04,A61K39/395,A61K39/
            PC   395,A61P31/06,
            PC   (C12N1/21,C12R1:19),C12N15/00,C12N5/00,A61K37/02 CC
            Compounds and methods for diagnosis of tuberculosis. FH   Key
                 Location/Qualifiers
            FT   source          1. .447
            FT                   /organism='Mycobacterium tuberculosis'.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Mycobacterium tuberculosis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1773"
ORIGIN

Alignment Scores:
Pred. No.:                 1.18e-13        Length:          447
Score:                     148.00          Matches:         30
Percent Similarity:        100.00%         Conservative:    0
Best Local Similarity:     100.00%         Mismatches:      0
Query Match:               100.00%         Indels:          0
DB:                        6               Gaps:            0

US-09-684-215B-17 (1-30) x BD205817 (1-447)

Qy           1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy          21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
               ||||||||||||||||||||||||||||||
Db          71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 13
AR261272
LOCUS       AR261272                675 bp    DNA     linear   PAT 29-JAN-2003
DEFINITION  Sequence 822 from patent US 6321716.
ACCESSION   AR261272
VERSION     AR261272.1  GI:28072035
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Mashiki,Z. and Harada,J.
  TITLE     Negative pressure control apparatus for engine mounted in vehicle
  JOURNAL   Patent: US 6321716-A 822 27-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="unknown"
                     /mol_type="genomic DNA"
```

ORIGIN

Alignment Scores:

Pred. No.:	1.83e-13	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x AR261272 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 14
AR400535
LOCUS AR400535 675 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 822 from patent US 6620922.
ACCESSION AR400535
VERSION AR400535.1 GI:40144000
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 675)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H., Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W., Hepler,W.T. and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: US 6620922-A 822 16-SEP-2003;
FEATURES Location/Qualifiers
source 1. .675
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:

Pred. No.:	1.83e-13	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x AR400535 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 15
AR405802
LOCUS AR405802 675 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 822 from patent US 6630305.
ACCESSION AR405802
VERSION AR405802.1 GI:40154639
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 675)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H., Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W., Hepler,W.T. and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: US 6630305-A 822 07-OCT-2003;
FEATURES Location/Qualifiers
source 1. .675
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:

Pred. No.:	1.83e-13	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x AR405802 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 16
AX201049
LOCUS AX201049 675 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 679 from Patent WO0151633.
ACCESSION AX201049
VERSION AX201049.1 GI:15390857
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y., Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W., Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0151633-A 679 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .675
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:

Pred. No.:	1.83e-13	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x AX201049 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
```

```
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 17
AX267848
LOCUS       AX267848                675 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 822 from Patent WO0173032.
ACCESSION   AX267848
VERSION     AX267848.1  GI:16516494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 822 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.83e-13        Length:         675
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x AX267848 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 18
BD251334
LOCUS       BD251334                702 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof.
ACCESSION   BD251334
VERSION     BD251334.1  GI:33061104
KEYWORDS    JP 2002510494-A/13.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Skeiky,Y.A.W., Alderson,M. and Neto,A.C.
  TITLE     Fused protein of Mycobacterium tuberculosis antigen and utilization
            thereof
  JOURNAL   Patent: JP 2002510494-A 13 09-APR-2002;
            CORIXA CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002510494-A/13
            PD   09-APR-2002
            PF   07-APR-1999 JP 2000542460
            PR   07-APR-1998 US    09/056556,30-DEC-1998 US    09/223040 PI
            YASIR A W SKEIKY,MARK ALDERSON,ANTONIO CAMPOS NETO PC
            C12N15/09,A61K39/04,A61K48/00,A61P31/04,C07K14/35,C07K19/00, PC
            C12P21/02,
            PC   C12N15/00
            CC   Description of Artificial Sequence:bi-fusion
            protein Ra12-DPPD
            CC          (designated
            CC   Mtb24), reading frame 1
            CC   bi-fusion protein Ra12-DPPD (designated Mtb24),reading frame 1
            CC   reading frame 2
            CC   reading frame 3
            FH   Key             Location/Qualifiers
            FT   CDS             (1). .(696)
            FT   CDS             (2). .(700)
            FT   CDS             (3). .(701).
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:              1.91e-13        Length:         702
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x BD251334 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114

RESULT 19
AR403747
LOCUS       AR403747                702 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 27 from patent US 6627198.
ACCESSION   AR403747
VERSION     AR403747.1  GI:40151423
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Reed,S.G., Skeiky,Y.A., Dillon,D.C., Alderson,M. and Campos-Neto,A.
  TITLE     Fusion proteins of Mycobacterium tuberculosis antigens and their
            uses
  JOURNAL   Patent: US 6627198-A 27 30-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .702
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.91e-13        Length:         702
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x AR403747 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
```

```
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114

RESULT 20
AR277645
LOCUS       AR277645                822 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 1862 from patent US 6509448.
ACCESSION   AR277645
VERSION     AR277645.1  GI:29711294
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 822)
  AUTHORS   Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
            Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6509448-A 1862 21-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                 2.26e-13       Length:          822
Score:                     148.00         Matches:         30
Percent Similarity:        100.00%        Conservative:    0
Best Local Similarity:     100.00%        Mismatches:      0
Query Match:               100.00%        Indels:          0
DB:                        6              Gaps:            0

US-09-684-215B-17 (1-30) x AR277645 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 21
AX369152
LOCUS       AX369152                822 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 1862 from Patent WO0204514.
ACCESSION   AX369152
VERSION     AX369152.1  GI:18857170
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
            Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
            Mcnabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0204514-A 1862 17-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                 2.26e-13       Length:          822
Score:                     148.00         Matches:         30
Percent Similarity:        100.00%        Conservative:    0
Best Local Similarity:     100.00%        Mismatches       0
Query Match:               100.00%        Indels:          0
DB:                        6              Gaps:            0

US-09-684-215B-17 (1-30) x AX369152 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 22
AX369167
LOCUS       AX369167                861 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 1877 from Patent WO0204514.
ACCESSION   AX369167
VERSION     AX369167.1  GI:18857177
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,
            Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S.,
            Mcnabb,A., Fanger,N., Switzer,A., Mcneill,P.D. and Clapper,J.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0204514-A 1877 17-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .861
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                 2.38e-13       Length:          861
Score:                     148.00         Matches:         30
Percent Similarity:        100.00%        Conservative:    0
Best Local Similarity:     100.00%        Mismatches:      0
Query Match:               100.00%        Indels:          0
DB:                        6              Gaps:            0

US-09-684-215B-17 (1-30) x AX369167 (1-861)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 23
AX351489
LOCUS       AX351489                894 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 236 from Patent WO0196390.
ACCESSION   AX351489
VERSION     AX351489.1  GI:18616835
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Hepler,W.T., Clapper,J.D., Wang,A. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196390-A 236 20-DEC-2001;
```

```
          CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.48e-13        Length:         894
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x AX351489 (1-894)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 24
AR220690
LOCUS       AR220690                900 bp    DNA     linear   PAT 26-SEP-2002
DEFINITION  Sequence 353 from patent US 6426072.
ACCESSION   AR220690
VERSION     AR220690.1  GI:23327471
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 900)
  AUTHORS   Wang,T., Fan,L., Kalos,M.D., Bangur,C.S., Hosken,N.A., Fanger,G.R.,
            Li,S.X., Wang,A., Skeiky,Y.A.W., Henderson,R.A. and McNeill,P.D.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: US 6426072-A 353 30-JUL-2002;
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.49e-13        Length:         900
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x AR220690 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 25
AX365960
LOCUS       AX365960                900 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 353 from Patent WO0200174.
ACCESSION   AX365960
VERSION     AX365960.1  GI:18697458
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Wang,T., Wang,A., Skeiky,Y.A., Li,S.X., Kalos,M.D., Henderson,R.A.,
            Mcneill,P.D., Fanger,N., Retter,M.W., Marnerakis,M., Fanger,G.R.,
            Vedvick,T.S., Carter,D., Watanabe,Y. and Peckham,D.W.
  TITLE     Compositions and methods for the therapy and diagnosis of lung
            cancer
  JOURNAL   Patent: WO 0200174-A 353 03-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .900
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.49e-13        Length:         900
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x AX365960 (1-900)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 26
AR261281
LOCUS       AR261281                915 bp    DNA     linear   PAT 29-JAN-2003
DEFINITION  Sequence 834 from patent US 6321716.
ACCESSION   AR261281
VERSION     AR261281.1  GI:28072044
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Mashiki,Z. and Harada,J.
  TITLE     Negative pressure control apparatus for engine mounted in vehicle
  JOURNAL   Patent: US 6321716-A 834 27-NOV-2001;
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              2.54e-13        Length:         915
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x AR261281 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
```

```
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 27
AR400544
LOCUS       AR400544                915 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 834 from patent US 6620922.
ACCESSION   AR400544
VERSION     AR400544.1  GI:40144016
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6620922-A 834 16-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                 2.54e-13      Length:          915
Score:                     148.00        Matches:         30
Percent Similarity:        100.00%       Conservative:    0
Best Local Similarity:     100.00%       Mismatches:      0
Query Match:               100.00%       Indels:          0
DB:                        6             Gaps:            0

US-09-684-215B-17 (1-30) x AR400544 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 28
AR405811
LOCUS       AR405811                915 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 834 from patent US 6630305.
ACCESSION   AR405811
VERSION     AR405811.1  GI:40154648
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 915)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
            Hepler,W.T. and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: US 6630305-A 834 07-OCT-2003;
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                 2.54e-13      Length:          915
Score:                     148.00        Matches:         30
Percent Similarity:        100.00%       Conservative:    0
Best Local Similarity:     100.00%       Mismatches:      0
Query Match:               100.00%       Indels:          0
DB:                        6             Gaps:            0

US-09-684-215B-17 (1-30) x AR405811 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 29
AX201061
LOCUS       AX201061                915 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 691 from Patent WO0151633.
ACCESSION   AX201061
VERSION     AX201061.1  GI:15390868
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 691 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .915
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                 2.54e-13      Length:          915
Score:                     148.00        Matches:         30
Percent Similarity:        100.00%       Conservative:    0
Best Local Similarity:     100.00%       Mismatches:      0
Query Match:               100.00%       Indels:          0
DB:                        6             Gaps:            0

US-09-684-215B-17 (1-30) x AX201061 (1-915)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 30
AX267860
LOCUS       AX267860                915 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 834 from Patent WO0173032.
ACCESSION   AX267860
VERSION     AX267860.1  GI:16516503
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
```

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0173032-A 834 04-OCT-2001;
CORIXA CORPORATION (US)

FEATURES Location/Qualifiers
source 1. .915
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	2.54e-13	Length:	915
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x AX267860 (1-915)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

Search completed: April 30, 2004, 02:24:01
Job time : 861.167 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 22:25:27 ; Search Time 587.584 Seconds
(without alignments)
1524.660 Million cell updates/sec

Title: US-09-684-215B-17
Perfect score: 148
Sequence: 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061305_13200/app_query.fasta_1.1180
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_5167_@runat_29042004_061305_13200 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	60	40.5	747	14	CB597827	CB597827 AGENCOURT
2	58	39.2	620	14	CF231012	CF231012 PtaC0016C
c 3	58	39.2	899	14	CF619317	CF619317 AGENCOURT
c 4	57	38.5	238	10	BB077789	BB077789 BB077789
5	57	38.5	336	14	CB184974	CB184974 PopSC0010
6	57	38.5	343	14	CA930964	CA930964 MTU2TA.P1
7	57	38.5	351	12	BI070561	BI070561 C040P04U
8	57	38.5	366	12	BI055877	BI055877 CM3-GN033
c 9	57	38.5	370	14	CA931030	CA931030 MTU2TA.P1
10	57	38.5	370	14	CA931960	CA931960 MTU4TA.P2
11	57	38.5	370	14	CA931997	CA931997 MTU4TA.P2
12	57	38.5	370	14	CA932035	CA932035 MTU4TA.P2
13	57	38.5	370	14	CA932108	CA932108 MTU4TA.P2
14	57	38.5	370	14	CA932138	CA932138 MTU4TA.P2
c 15	57	38.5	374	14	CA927770	CA927770 MTU6TR.P1
c 16	57	38.5	374	14	CA928449	CA928449 MTU6TR.P2
c 17	57	38.5	374	14	CA928493	CA928493 MTU6TR.P3
c 18	57	38.5	374	14	CA928769	CA928769 MTU6TR.P6
c 19	57	38.5	384	14	CA931116	CA931116 MTU2TA.P2
c 20	57	38.5	385	14	CA931394	CA931394 MTU2TA.P6
c 21	57	38.5	385	14	CA931501	CA931501 MTU2TA.P8
22	57	38.5	387	14	CA931767	CA931767 MTU4TA.P2
23	57	38.5	388	14	CA927636	CA927636 MTU6TR.P1
c 24	57	38.5	388	14	CA928025	CA928025 MTU6TR.P1
c 25	57	38.5	388	14	CA929028	CA929028 MTU6TR.P9
c 26	57	38.5	389	14	CA927727	CA927727 MTU6TR.P1
c 27	57	38.5	389	14	CA928142	CA928142 MTU6TR.P1
28	57	38.5	390	14	CA927617	CA927617 MTU6TR.P1
c 29	57	38.5	390	14	CA927744	CA927744 MTU6TR.P1
c 30	57	38.5	390	14	CA927930	CA927930 MTU6TR.P1
c 31	57	38.5	402	14	CA933902	CA933902 MTU3TS.P1
32	57	38.5	403	13	BU824075	BU824075 UB60BPA07
c 33	57	38.5	404	14	CA934515	CA934515 MTU3TS.P8
34	57	38.5	405	14	CK099325	CK099325 A062P49.5
35	57	38.5	406	14	CF231169	CF231169 PtaC0018C
c 36	57	38.5	408	14	CA930794	CA930794 MTU2TA.P1
c 37	57	38.5	409	14	CA931302	CA931302 MTU2TA.P5
c 38	57	38.5	409	14	CA931312	CA931312 MTU2TA.P5
c 39	57	38.5	413	14	CA931446	CA931446 MTU2TA.P7
c 40	57	38.5	413	14	CA931603	CA931603 MTU2TA.P9
c 41	57	38.5	414	14	CA927769	CA927769 MTU6TR.P1
c 42	57	38.5	414	14	CA928453	CA928453 MTU6TR.P2
c 43	57	38.5	414	14	CA929036	CA929036 MTU6TR.P9
44	57	38.5	415	13	BU819016	BU819016 UA38BPC06
45	57	38.5	416	9	AI164854	AI164854 A069p72u

ALIGNMENTS

RESULT 1
CB597827/c
LOCUS CB597827 747 bp mRNA linear EST 03-APR-2003
DEFINITION AGENCOURT_12972312 NIH_MGC_178 Mus musculus cDNA clone IMAGE:30297156 5', mRNA sequence.
ACCESSION CB597827
VERSION CB597827.1 GI:29515683
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 747)

```
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM57  row: j  column: 13
            High quality sequence stop: 475.
FEATURES             Location/Qualifiers
     source          1. .747
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30297156"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_178"
                     /note="Organ: lung and heart; Vector: pDNR-LIB; Site_1:
                     SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); cDNA
                     made by oligo-dT priming and directionally cloned. 5' and
                     3' adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.5 kb
                     size fraction. Library created in the laboratory of M.
                     Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:                111          Length:          747
Score:                    60.00        Matches:         13
Percent Similarity:       70.83%       Conservative:    4
Best Local Similarity:    54.17%       Mismatches:      5
Query Match:              40.54%       Indels:          2
DB:                       14           Gaps:            1

US-09-684-215B-17 (1-30) x CB597827 (1-747)

Qy          9 LeuSerGlnGlyGly------GlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIle 26
              |||   |||||||||      :::|||||||     |||:::|||||||||:::
Db        635 TTGGGCCAGGGTGGGCATAGGCGAGGGTTCATTGAGCCCTTAGGTCAGGCAGTCACCACA 576

Qy         27 AlaGlyGlnIle 30
              ||||||||||:::
Db        575 GCGGGGCAGCTT 564

RESULT 2
CF231012
LOCUS       CF231012                620 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  PtaC0016C2C0206 Poplar cDNA library from cambial zone Populus alba
            x Populus tremula cDNA 5', mRNA sequence.
ACCESSION   CF231012
VERSION     CF231012.1  GI:33450441
KEYWORDS    EST.
SOURCE      Populus alba x Populus tremula
  ORGANISM  Populus alba x Populus tremula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 620)
  AUTHORS   Dejardin,A., Leple,J.-C., Lesage-Descauses,M.-C., Costa,G. and
            Pilate,G.
  TITLE     Expressed sequence tags from poplar tension wood tissues - A
            comparative analysis from multiple libraries
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Leple JC
            Unit of Forest improvement, Genetics and Physiology
            National Institute for Agricultural Research (INRA)
            Domaine de Limere, BP20619 ARDON, 45166 OLIVET CEDEX, FRANCE
            Tel: 33 02 38 41 78 00
            Fax: 33 02 38 41 78 79
            Email: Jean-Charles.Leple@orleans.inra.fr
            PCR PRimers
            FORWARD: TriplexA 5' CTCGGGAAGCGCGCCATTGTG 3'
            BACKWARD: TriplexB1 5' ATACGACTCACTATAGGGCGA 3'
            Plate: PtaC0016  row: C  column: 2
            Seq primer: TriplexA 5' CTCGGGAAGCGCGCCATTGTG 3'.
FEATURES             Location/Qualifiers
     source          1. .620
                     /organism="Populus alba x Populus tremula"
                     /mol_type="mRNA"
                     /strain="clone INRA 717-1-B4"
                     /db_xref="taxon:80863"
                     /sex="female"
                     /tissue_type="cambial zone harvested on the bark side"
                     /dev_stage="3-years-old poplar trees grown in the nursery"
                     /clone_lib="Poplar cDNA library from cambial zone"
                     /note="A composite cDNA library was made with mRNA
                     isolated from opposite and tension wood tissues
                     corresponding to the cambial zone collected on the bark
                     side after debarking the stem. In this respect, in
                     addition to cambium cDNA, this library also contains very
                     young phloem and very young xylem cDNA. The sampling was
                     done on 3 different tilted trees grown in the nursery.
                     cDNA were cloned in an oriented way into SfiI (A and B)
                     restriction sites. A one-step conversion of Lambda
                     TriplEx2 to the corresponding pTriplEx2 plasmid was done
                     via site-specific recombination at loxP sites (Clontech ;
                     SMART cDNA library construction kit). cDNA inserts were
                     PCR amplified using flanking primers and then sequenced on
                     a ABI3100 Genetic Analyser (Applied Biosystem)"
ORIGIN

Alignment Scores:
Pred. No.:                174          Length:          620
Score:                    58.00        Matches:         10
Percent Similarity:       66.67%       Conservative:    6
Best Local Similarity:    41.67%       Mismatches:      8
Query Match:              39.19%       Indels:          0
DB:                       14           Gaps:            0

US-09-684-215B-17 (1-30) x CF231012 (1-620)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||:::::   |||||||||   |||||||||
Db        261 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 320

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        321 GGAGAAGTGATG 332

RESULT 3
CF619317/c
LOCUS       CF619317                899 bp    mRNA    linear   EST 01-OCT-2003
DEFINITION  AGENCOURT_15738624 NIH_MGC_219 Homo sapiens cDNA clone
            IMAGE:30523641 5', mRNA sequence.
ACCESSION   CF619317
VERSION     CF619317.1  GI:37238006
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 899)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
```

Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM601 row: o column: 10
High quality sequence start: 55
High quality sequence stop: 553.

FEATURES Location/Qualifiers
source 1. .899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30523641"
/tissue_type="Pooled Chondrosarcoma Tumor cells"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_219"
/note="Vector: pYX-Asc; Site_1: EcoRI; Site_2: NotI; Library is oligo-dT primed and directionally cloned Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according tomRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 0.5-1Kb. Adaptors 5'(AATTCGGCACGAGG)3' and 5'd (CCTCGTGCCG)3'. 3' Linker sequence - GCGGCCGCTGAGAGCC T18. Sequencing primers 3'end: T3 promoter primer 5'd (ATTAACCCTCACTAAAGGGA)3'. 5' End: T7 promoter primer 5'd (TAATACGACTCACTATAGGG)3'. Library was constructed in the laboratory of M. Bento Soares. Average insert size 2-3kb. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	281	Length:	899
Score:	58.00	Matches:	11
Percent Similarity:	61.54%	Conservative:	5
Best Local Similarity:	42.31%	Mismatches:	10
Query Match:	39.19%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-17 (1-30) x CF619317 (1-899)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||                  |||      |||:::|||||||||||||   |||:::|||
Db        677 ACGGGGGTTGGGTAGGGGTTCAACTTCTCAAAGGGCGGGCAGGGGGCAGCTCTTCCTCCG 618

Qy         21 GlyGlnAlaMetAlaIle 26
              |||:::   :::|||:::
Db        617 GGGAGGCCCCTGGCAGTA 600
```

RESULT 4
BB077789/c
LOCUS BB077789 238 bp mRNA linear EST 27-JUN-2000
DEFINITION BB077789 RIKEN full-length enriched, adult male diencephalon Mus musculus cDNA clone 9330145E04 3' similar to AF178432 Homo sapiens SH3 protein (AF3P21) mRNA, mRNA sequence.
ACCESSION BB077789
VERSION BB077789.1 GI:8642849
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 238)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .238
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9330145E04"
/sex="male"
/tissue_type="diencephalon"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male diencephalon"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

ORIGIN

Alignment Scores:

Pred. No.:	71	Length:	238
Score:	57.00	Matches:	9
Percent Similarity:	75.00%	Conservative:	6

```
Best Local Similarity:  45.00%          Mismatches:     5
Query Match:            38.51%          Indels:         0
DB:                     10              Gaps:           0

US-09-684-215B-17 (1-30) x BB077789 (1-238)

Qy          9 LeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGlnAlaMetAlaIleAlaGly 28
              |||      |||||||||||||   :::::::|||:::|||:::   :::   |||:::|||
Db        112 CTGAAAGGAGGGGGTCAGGGGATATCAGTCCCAATGGGGGAAGGGCTGCAAATTAGCGGG 53

RESULT 5
CB184974
LOCUS       CB184974                336 bp    mRNA    linear   EST 03-FEB-2003
DEFINITION  PopSC00108 Poplar SC cDNA library Populus alba x Populus glandulosa
            cDNA clone PopSC00108, mRNA sequence.
ACCESSION   CB184974
VERSION     CB184974.1  GI:28196969
KEYWORDS    EST.
SOURCE      Populus alba x Populus glandulosa
  ORGANISM  Populus alba x Populus glandulosa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 336)
  AUTHORS   Lee,J.S., Lee,H.S., Noh,E.W. and Choi,Y.I.
  TITLE     Gene Expression Profiling of the Poplar Suspension Cell
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Jae-Soon Lee
            Biotechnology Team
            Korea Forest Research Institute
            44-3 Omockchun-dong, Suwon, 441-350, Korea
            Tel: 82 31 290 1162
            Fax: 82 31 290 1020
            Email: jasolee@foa.go.kr
            PCR PRimers
            FORWARD: T3
            BACKWARD: T7
            Insert Length: 336   Std Error: 0.00
            Seq primer: T3
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .336
                     /organism="Populus alba x Populus glandulosa"
                     /mol_type="mRNA"
                     /db_xref="taxon:153471"
                     /clone="PopSC00108"
                     /tissue_type="Suspension cell"
                     /lab_host="XL1-Blue MRF' strain"
                     /clone_lib="Poplar SC cDNA library"
                     /note="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI"
ORIGIN

Alignment Scores:
Pred. No.:              111             Length:         336
Score:                  57.00           Matches:        10
Percent Similarity:     66.67%          Conservative:   6
Best Local Similarity:  41.67%          Mismatches:     8
Query Match:            38.51%          Indels:         0
DB:                     14              Gaps:           0

US-09-684-215B-17 (1-30) x CB184974 (1-336)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        243 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 302

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        303 GGAGAAATGATG 314

RESULT 6
CA930964
LOCUS       CA930964                343 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU2TA.P13.A01 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA930964
VERSION     CA930964.1  GI:27419444
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 343)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .343
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:              114             Length:         343
Score:                  57.00           Matches:        10
Percent Similarity:     66.67%          Conservative:   6
Best Local Similarity:  41.67%          Mismatches:     8
Query Match:            38.51%          Indels:         0
DB:                     14              Gaps:           0

US-09-684-215B-17 (1-30) x CA930964 (1-343)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        198 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 257

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        258 GGAGAAATGATG 269

RESULT 7
BI070561
LOCUS       BI070561                351 bp    mRNA    linear   EST 19-JUN-2001
DEFINITION  C040P04U Populus strain T89 leaves Populus tremula x Populus
            tremuloides cDNA, mRNA sequence.
ACCESSION   BI070561
VERSION     BI070561.1  GI:14491181
KEYWORDS    EST.
SOURCE      Populus tremula x Populus tremuloides
  ORGANISM  Populus tremula x Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 351)
  AUTHORS   Jansson,S., Bhalerao,R., Erlandsson,R., Bjorkbacka,H., Karlsson,J.,
            Sterky,F., Gustafsson,P. and Lundeberg,J.
  TITLE     Gene expression in Populus leaves
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Erlandsson R
            Department of Biotechnology
            Royal Institute of Technology
            Teknikringen 30, Stockholm S-10044, Sweden
            Tel: 46 8 790 8287
```

```
            Fax: 46 8 245452
            Email: rikerl@biochem.kth.se.
FEATURES              Location/Qualifiers
     source           1. .351
                      /organism="Populus tremula x Populus tremuloides"
                      /mol_type="mRNA"
                      /strain="T89"
                      /db_xref="taxon:47664"
                      /tissue_type="leaf"
                      /clone_lib="Populus strain T89 leaves"
ORIGIN

Alignment Scores:
Pred. No.:              117           Length:         351
Score:                  57.00         Matches:        10
Percent Similarity:     66.67%        Conservative:   6
Best Local Similarity:  41.67%        Mismatches:     8
Query Match:            38.51%        Indels:         0
DB:                     12            Gaps:           0

US-09-684-215B-17 (1-30) x BI070561 (1-351)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||:::::::   |||||||||   |||||||||
Db         32 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 91

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         92 GGAGAAATGATG 103

RESULT 8
BI055877
LOCUS       BI055877                366 bp    mRNA    linear   EST 15-JUN-2001
DEFINITION  CM3-GN0330-120201-735-e08 GN0330 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BI055877
VERSION     BI055877.1  GI:14463407
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM3&t2=CM3-GN0330-
            120201-735-e08&t3=2001-02-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 366.
FEATURES              Location/Qualifiers
     source           1. .366
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="GN0330"
                      /note="Organ: placenta_normal; Vector: puc18; Site_1:
                      SmaI; Site_2: SmaI; A mini-library was made by cloning
                      products derived from  ORESTES PCR (U.S. Letters Patent
                      application No. 196,716 -  Ludwig Institute for Cancer
                      Research) profiles into  the pUC 18 vector. Reverse
                      transcription of tissue mRNA and cDNA amplification were
                      performed under low stringency conditions."
ORIGIN

Alignment Scores:
Pred. No.:              124           Length:         366
Score:                  57.00         Matches:        10
Percent Similarity:     71.43%        Conservative:   5
Best Local Similarity:  47.62%        Mismatches:     6
Query Match:            38.51%        Indels:         0
DB:                     12            Gaps:           0

US-09-684-215B-17 (1-30) x BI055877 (1-366)

Qy          3 AlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIleGlyGln 22
              |||   :::     ||||||   :::   |||||||||   |||:::|||:::|||:::
Db        184 GCACCGGAGGAGGCACAGCTCGATCGCCCTGGGCAAGGGATTGCTGTGCCTGTGGGCGAG 243

Qy         23 Ala 23
              |||
Db        244 GCC 246

RESULT 9
CA931030/c
LOCUS       CA931030                370 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU2TA.P1.A09 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA931030
VERSION     CA931030.1  GI:27419510
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES              Location/Qualifiers
     source           1. .370
                      /organism="Populus tremuloides"
                      /mol_type="mRNA"
                      /db_xref="taxon:3693"
                      /clone_lib="Aspen apex cDNA Library"
                      /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:              126           Length:         370
Score:                  57.00         Matches:        10
Percent Similarity:     66.67%        Conservative:   6
Best Local Similarity:  41.67%        Mismatches:     8
Query Match:            38.51%        Indels:         0
DB:                     14            Gaps:           0

US-09-684-215B-17 (1-30) x CA931030 (1-370)
```

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 10
CA931960
LOCUS       CA931960                370 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU4TA.P25.B07 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA931960
VERSION     CA931960.1  GI:27420440
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:              126             Length:         370
Score:                  57.00           Matches:        10
Percent Similarity:     66.67%          Conservative:   6
Best Local Similarity:  41.67%          Mismatches:     8
Query Match:            38.51%          Indels:         0
DB:                     14              Gaps:           0

US-09-684-215B-17 (1-30) x CA931960 (1-370)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        225 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 284

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        285 GGAGAAATGATG 296

RESULT 11
CA931997
LOCUS       CA931997                370 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU4TA.P25.F03 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA931997
VERSION     CA931997.1  GI:27420477
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:              126             Length:         370
Score:                  57.00           Matches:        10
Percent Similarity:     66.67%          Conservative:   6
Best Local Similarity:  41.67%          Mismatches:     8
Query Match:            38.51%          Indels:         0
DB:                     14              Gaps:           0

US-09-684-215B-17 (1-30) x CA931997 (1-370)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        225 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 284

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        285 GGAGAAATGATG 296

RESULT 12
CA932035
LOCUS       CA932035                370 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU4TA.P26.A11 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA932035
VERSION     CA932035.1  GI:27420515
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
```

```
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:                126             Length:         370
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA932035 (1-370)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        225 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 284

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        285 GGAGAAATGATG 296

RESULT 13
CA932108
LOCUS       CA932108                370 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU4TA.P27.A02 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA932108
VERSION     CA932108.1  GI:27420588
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:                126             Length:         370
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA932108 (1-370)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        225 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 284

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        285 GGAGAAATGATG 296

RESULT 14
CA932138
LOCUS       CA932138                370 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU4TA.P27.D05 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA932138
VERSION     CA932138.1  GI:27420618
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:                126             Length:         370
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA932138 (1-370)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        225 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 284

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        285 GGAGAAATGATG 296

RESULT 15
CA927770/c
LOCUS       CA927770                374 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P12.G06 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA927770
VERSION     CA927770.1  GI:27416249
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 374)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
```

Plant Biotech Research Center
Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

FEATURES Location/Qualifiers
source 1. .374
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen root cDNA Library"
/note="Organ: root"

ORIGIN

Alignment Scores:

Pred. No.:	127	Length:	374
Score:	57.00	Matches:	10
Percent Similarity:	66.67%	Conservative:	6
Best Local Similarity:	41.67%	Mismatches:	8
Query Match:	38.51%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-17 (1-30) x CA927770 (1-374)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75
```

RESULT 16
CA928449/c
LOCUS CA928449 374 bp mRNA linear EST 30-DEC-2002
DEFINITION MTU6TR.P2.G12 Aspen root cDNA Library Populus tremuloides cDNA, mRNA sequence.
ACCESSION CA928449
VERSION CA928449.1 GI:27416928
KEYWORDS EST.
SOURCE Populus tremuloides (quaking aspen)
ORGANISM Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE 1 (bases 1 to 374)
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai,C-J.
TITLE Expressed sequence tags from Aspen
JOURNAL Unpublished (2003)
COMMENT Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources & Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.
FEATURES Location/Qualifiers
source 1. .374
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen root cDNA Library"
/note="Organ: root"
ORIGIN

Alignment Scores:

Pred. No.:	127	Length:	374
Score:	57.00	Matches:	10
Percent Similarity:	66.67%	Conservative:	6
Best Local Similarity:	41.67%	Mismatches:	8
Query Match:	38.51%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-17 (1-30) x CA928449 (1-374)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGGGAAATGATG 75
```

RESULT 17
CA928493/c
LOCUS CA928493 374 bp mRNA linear EST 30-DEC-2002
DEFINITION MTU6TR.P3.C12 Aspen root cDNA Library Populus tremuloides cDNA, mRNA sequence.
ACCESSION CA928493
VERSION CA928493.1 GI:27416972
KEYWORDS EST.
SOURCE Populus tremuloides (quaking aspen)
ORGANISM Populus tremuloides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE 1 (bases 1 to 374)
AUTHORS Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and Tsai,C-J.
TITLE Expressed sequence tags from Aspen
JOURNAL Unpublished (2003)
COMMENT Contact: Tsai C-J
Plant Biotech Research Center
Michigan Technological University, School of Forest Resources & Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.
FEATURES Location/Qualifiers
source 1. .374
/organism="Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:3693"
/clone_lib="Aspen root cDNA Library"
/note="Organ: root"
ORIGIN

Alignment Scores:

Pred. No.:	127	Length:	374
Score:	57.00	Matches:	10
Percent Similarity:	66.67%	Conservative:	6
Best Local Similarity:	41.67%	Mismatches:	8
Query Match:	38.51%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-17 (1-30) x CA928493 (1-374)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75
```

RESULT 18
CA928769/c
LOCUS CA928769 374 bp mRNA linear EST 30-DEC-2002
DEFINITION MTU6TR.P6.F02 Aspen root cDNA Library Populus tremuloides cDNA, mRNA sequence.
ACCESSION CA928769

```
VERSION     CA928769.1  GI:27417250
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 374)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .374
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN

Alignment Scores:
Pred. No.:                127             Length:         374
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA928769 (1-374)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 19
CA931116/c
LOCUS       CA931116                384 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU2TA.P2.D03 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA931116
VERSION     CA931116.1  GI:27419596
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 384)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .384
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:                132             Length:         384
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA931116 (1-384)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 20
CA931394/c
LOCUS       CA931394                385 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU2TA.P6.G03 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA931394
VERSION     CA931394.1  GI:27419874
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 385)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .385
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:                132             Length:         385
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA931394 (1-385)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
```

```
              |||:::   ::::::   |||:::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 21
CA931501/c
LOCUS       CA931501                385 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU2TA.P8.B06 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA931501
VERSION     CA931501.1  GI:27419981
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 385)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .385
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:                 132           Length:          385
Score:                     57.00         Matches:         10
Percent Similarity:        66.67%        Conservative:    6
Best Local Similarity:     41.67%        Mismatches:      8
Query Match:               38.51%        Indels:          0
DB:                        14            Gaps:            0

US-09-684-215B-17 (1-30) x CA931501 (1-385)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||:::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 22
CA931767
LOCUS       CA931767                387 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU4TA.P22.F11 Aspen apex cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA931767
VERSION     CA931767.1  GI:27420247
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 387)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .387
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen apex cDNA Library"
                     /note="Organ: apex"
ORIGIN

Alignment Scores:
Pred. No.:                 133           Length:          387
Score:                     57.00         Matches:         10
Percent Similarity:        66.67%        Conservative:    6
Best Local Similarity:     41.67%        Mismatches:      8
Query Match:               38.51%        Indels:          0
DB:                        14            Gaps:            0

US-09-684-215B-17 (1-30) x CA931767 (1-387)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||:::::::   |||||||||   |||||||||
Db        242 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 301

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        302 GGAGAAATGATG 313

RESULT 23
CA927636
LOCUS       CA927636                388 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P11.C07 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA927636
VERSION     CA927636.1  GI:27416115
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .388
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
```

```
ORIGIN

Alignment Scores:
Pred. No.:                134             Length:         388
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA927636 (1-388)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        243 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 302

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        303 GGAGAAATGATG 314

RESULT 24
CA928025/c
LOCUS       CA928025                388 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P15.H06 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA928025
VERSION     CA928025.1  GI:27416504
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .388
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN

Alignment Scores:
Pred. No.:                134             Length:         388
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA928025 (1-388)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 25
CA929028/c
LOCUS       CA929028                388 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P9.F10 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA929028
VERSION     CA929028.1  GI:27417509
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 388)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .388
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN

Alignment Scores:
Pred. No.:                134             Length:         388
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA929028 (1-388)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCGGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 26
CA927727/c
LOCUS       CA927727                389 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P12.C07 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA927727
VERSION     CA927727.1  GI:27416206
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
```

Michigan Technological University, School of Forest Resources &
Environmental Science
1400 Townsend Drive, Houghton, MI 49931-1295, USA
Tel: 906 487 2914
Fax: 906 487 2915
Email: chtsai@mtu.edu.

```
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN
```

Alignment Scores:

Pred. No.:	134	Length:	389
Score:	57.00	Matches:	10
Percent Similarity:	66.67%	Conservative:	6
Best Local Similarity:	41.67%	Mismatches:	8
Query Match:	38.51%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-17 (1-30) x CA927727 (1-389)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75
```

```
RESULT 27
CA928142/c
LOCUS       CA928142                389 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P17.C03 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA928142
VERSION     CA928142.1  GI:27416621
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN
```

Alignment Scores:

Pred. No.:	134	Length:	389
Score:	57.00	Matches:	10
Percent Similarity:	66.67%	Conservative:	6
Best Local Similarity:	41.67%	Mismatches:	8
Query Match:	38.51%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-17 (1-30) x CA928142 (1-389)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCGGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75
```

```
RESULT 28
CA927617
LOCUS       CA927617                390 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P11.A10 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA927617
VERSION     CA927617.1  GI:27416096
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .390
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN
```

Alignment Scores:

Pred. No.:	134	Length:	390
Score:	57.00	Matches:	10
Percent Similarity:	66.67%	Conservative:	6
Best Local Similarity:	41.67%	Mismatches:	8
Query Match:	38.51%	Indels:	0
DB:	14	Gaps:	0

US-09-684-215B-17 (1-30) x CA927617 (1-390)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   |||||||||   |||||||||
Db        245 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 304

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db        305 GGAGAAATGATG 316
```

```
RESULT 29
CA927744/c
LOCUS       CA927744                390 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P12.E03 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA927744
VERSION     CA927744.1  GI:27416223
```

```
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .390
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN

Alignment Scores:
Pred. No.:                134             Length:         390
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA927744 (1-390)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   ||||||||   ||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75

RESULT 30
CA927930/c
LOCUS       CA927930                390 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  MTU6TR.P14.F11 Aspen root cDNA Library Populus tremuloides cDNA,
            mRNA sequence.
ACCESSION   CA927930
VERSION     CA927930.1  GI:27416409
KEYWORDS    EST.
SOURCE      Populus tremuloides (quaking aspen)
  ORGANISM  Populus tremuloides
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Ranjan,P., Kao,Y.-Y., Harding,S.A., Jiang,H., Joshi,C.P. and
            Tsai,C-J.
  TITLE     Expressed sequence tags from Aspen
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Tsai C-J
            Plant Biotech Research Center
            Michigan Technological University, School of Forest Resources &
            Environmental Science
            1400 Townsend Drive, Houghton, MI 49931-1295, USA
            Tel: 906 487 2914
            Fax: 906 487 2915
            Email: chtsai@mtu.edu.
FEATURES             Location/Qualifiers
     source          1. .390
                     /organism="Populus tremuloides"
                     /mol_type="mRNA"
                     /db_xref="taxon:3693"
                     /clone_lib="Aspen root cDNA Library"
                     /note="Organ: root"
ORIGIN

Alignment Scores:
Pred. No.:                134             Length:         390
Score:                    57.00           Matches:        10
Percent Similarity:       66.67%          Conservative:   6
Best Local Similarity:    41.67%          Mismatches:     8
Query Match:              38.51%          Indels:         0
DB:                       14              Gaps:           0

US-09-684-215B-17 (1-30) x CA927930 (1-390)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              |||:::   ::::::   |||::::::   ||||||||   ||||||||
Db        146 ACATCTGGAGCAGAGGGATTCGAAGTGTTCCAAGGAGGATCAGGGTTCGCATTTCTCGCA 87

Qy         21 GlyGlnAlaMet 24
              |||:::   |||
Db         86 GGAGAAATGATG 75


Search completed: April 30, 2004, 04:40:53
Job time : 590.584 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 29, 2004, 20:30:12 ; Search time 85.3349 Seconds
(without alignments)
1493.479 Million cell updates/sec

Title: US-09-684-215B-17
Perfect score: 148
Sequence: 1 TAASDNFQLSQGGQGFAIPIGQAMAIAGQI 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09684215/runat_29042004_061304_13179/app_query.fasta_1.1180
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=30
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09684215_@CGN_1_1_819_@runat_29042004_061304_13179 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	148	100.0	186	6	ABK39777	Abk39777 DNA encod
2	148	100.0	186	7	ACA12106	Aca12106 Human His
3	148	100.0	186	7	ACA03292	Aca03292 Lung canc
4	148	100.0	396	4	AAL40769	Aal40769 Nucleotid
5	148	100.0	399	6	ABS71801	Abs71801 Human DNA
6	148	100.0	399	9	AAD60769	Aad60769 Ra12 (s)-
7	148	100.0	447	2	AAT91403	Aat91403 Mycobacte
8	148	100.0	447	2	AAT91466	Aat91466 Mycobacte
9	148	100.0	447	2	AAV44342	Aav44342 Mycobacte
10	148	100.0	447	2	AAV64450	Aav64450 M. tuberc
11	148	100.0	447	2	AAZ19040	Aaz19040 M. tuberc
12	148	100.0	447	2	AAZ19252	Aaz19252 M. tuberc
13	148	100.0	447	5	AAS03780	Aas03780 M. tuberc
14	148	100.0	447	6	AAD47080	Aad47080 Mycobacte
15	148	100.0	447	6	AAD28339	Aad28339 Mycobacte
16	148	100.0	672	4	AAL40772	Aal40772 Nucleotid
17	148	100.0	675	4	AAH93896	Aah93896 Ra12-P510
18	148	100.0	675	4	AAS64132	Aas64132 Human /M.
19	148	100.0	675	5	ACA59940	Aca59940 Prostate
20	148	100.0	675	6	ABL95503	Abl95503 Ra12- P51
21	148	100.0	675	7	ACC95667	Acc95667 Prostate
22	148	100.0	675	9	ADB14272	Adb14272 Human pro
23	148	100.0	702	2	AAZ20206	Aaz20206 Mycobacte
24	148	100.0	702	4	AAL40770	Aal40770 Nucleotid
25	148	100.0	702	6	ABK14140	Abk14140 DNA encod
26	148	100.0	822	6	ABK39769	Abk39769 DNA encod
27	148	100.0	822	7	ACA12098	Aca12098 Human lun
28	148	100.0	822	7	ACA03284	Aca03284 Lung canc
29	148	100.0	861	6	ABK39776	Abk39776 DNA encod
30	148	100.0	861	7	ACA12105	Aca12105 Human Ra1
31	148	100.0	861	7	ACA03291	Aca03291 Lung canc
32	148	100.0	894	6	ABK27798	Abk27798 Human col
33	148	100.0	900	6	ABL49257	Abl49257 Ra12/C-te
34	148	100.0	900	6	ABQ92443	Abq92443 Human lun
35	148	100.0	900	8	ADA28443	Ada28443 Lung tumo
36	148	100.0	915	4	AAH93905	Aah93905 Ra12-P775
37	148	100.0	915	4	AAS64141	Aas64141 Human /M.
38	148	100.0	915	5	ACA59949	Aca59949 Prostate
39	148	100.0	915	6	ABL95512	Abl95512 Ra12- P77
40	148	100.0	915	7	ACC95676	Acc95676 Prostate
41	148	100.0	915	9	ADB14284	Adb14284 Human pro
42	148	100.0	945	6	ABK39768	Abk39768 DNA encod
43	148	100.0	945	7	ACA12097	Aca12097 Human lun
44	148	100.0	945	7	ACA03283	Aca03283 Lung canc
45	148	100.0	1002	6	AAD47078	Aad47078 Mycobacte

ALIGNMENTS

RESULT 1
ABK39777
ID ABK39777 standard; cDNA; 186 BP.
XX
AC ABK39777;
XX
DT 21-MAY-2002 (first entry)
XX
DE DNA encoding Ral2S-L985PEx peptide.
XX
KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW gene; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200204514-A2.
XX
PD 17-JAN-2002.
XX
PF 10-JUL-2001; 2001WO-US022058.
XX
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX
PA (CORI-) CORIXA CORP.

```
XX
PI   Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI   Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, Mcnabb A;
PI   Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
XX
DR   WPI; 2002-164634/21.
DR   P-PSDB; AAU85595.
XX
PT   Novel polynucleotide encoding a lung tumor polypeptide useful for
PT   stimulating and/or expanding T cells specific for a tumor protein.
XX
PS   Claim 1; SEQ ID NO 1879; 223pp; English.
XX
CC   The invention describes an isolated polynucleotide and polypeptide useful
CC   for stimulating and/or expanding T cells specific for a tumour protein
CC   for determining the presence of a cancer in a patient. A composition
CC   containing the polynucleotide and/or polypeptide is useful for treating a
CC   lung cancer in a patient. The polypeptide is useful for removing tumour
CC   cells from a biological sample. The polynucleotide is also useful as
CC   probe or primer to detect the level of mRNA encoding a tumour protein.
CC   This sequence encodes a lung tumour associated protein or protein
CC   fragment, described in the method of the invention. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.91e-15      Length:          186
Score:                    148.00        Matches:         30
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       6             Gaps:            0

US-09-684-215B-17 (1-30) x ABK39777 (1-186)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              |||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 2
ACA12106
ID   ACA12106 standard; cDNA; 186 BP.
XX
AC   ACA12106;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Human His-tagged Ra12S-L985PEx fusion protein cDNA.
XX
KW   Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW   T cell expansion; CD4; CD8; RA12; gene.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US2002197669-A1.
XX
PD   26-DEC-2002.
XX
PF   03-MAY-2001; 2001US-00849626.
XX
PR   13-DEC-2000; 2000US-00736457.
XX
PA   (BANG/) BANGUR C S.
PA   (FANG/) FANGER G R.
PA   (WANG/) WANG A.
PA   (WANG/) WANG T.
PA   (SWIT/) SWITZER A P.
PA   (MCNE/) MCNEILL P D.
PA   (CLAP/) CLAPPER J D.
XX
PI   Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;
PI   Clapper JD;
XX
DR   WPI; 2003-352750/33.
DR   P-PSDB; ABU69570.
XX
PT   Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT   detecting the presence of lung cancer in a patient, and in pharmaceutical
PT   compositions, e.g. vaccines, for treating lung cancer.
XX
PS   Example 10; Page; 72pp; English.
XX
CC   The invention relates to a polynucleotide encoding a lung tumour protein,
CC   comprising a sequence selected from any of the 14 sequences mentioned in
CC   the specification, or a sequence (S2) mentioned in specification,
CC   complement of S1, sequences consisting of at least 20 contiguous residues
CC   of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC   90%, identity to S1, or degenerate variants of S1. Also included are an
CC   isolated polypeptide (comprising a sequence (S3) selected from any one of
CC   the 4 amino acid sequences mentioned in the specification, a sequence
CC   encoded by the polynucleotide, or sequences having at least 70%,
CC   preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC   expression vector comprising the polynucleotide operably linked to an
CC   expression control sequence, a host cell transformed or transfected with
CC   the vector, an isolated antibody (or its antigen-binding fragment) that
CC   specifically binds to the polypeptide, detecting the presence of a cancer
CC   in a patient, a fusion protein comprising the polypeptide, an
CC   oligonucleotide that hybridises to S1 under moderately stringent
CC   conditions, stimulating and/or expanding T cells specific for a tumour
CC   protein (comprising contacting T cells with the polynucleotide, protein
CC   or antigen-presenting cells, under conditions and for a time sufficient
CC   to permit the stimulation and/or expansion of T cells) and inhibiting the
CC   development of a cancer in a patient (by incubating CD4^+ and/or CD8^+ T
CC   cells isolated from a patient with the polynucleotide, protein or antigen
CC   presenting cells that express the polynucleotide, such that T cells
CC   proliferate, administering to the patient an effective amount of the
CC   proliferated T cells, and thus inhibiting the development of a cancer in
CC   the patient. The polynucleotide, protein ands cells are useful in a
CC   composition for stimulating an immune response in a patient, and for
CC   treating a cancer in a patient (particularly lung cancer). The
CC   oligonucleotide is useful for determining the presence of a cancer in a
CC   patient. The protein and oligonucleotides are useful in pharmaceutical
CC   compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC   or primer for nucleic acid hybridisation, and in the design and
CC   preparation of ribozyme molecules for inhibiting expression of tumour
CC   polypeptides and proteins in tumour cells. An amplified portion of the
CC   polynucleotide is useful for isolating a full-length gene from a suitable
CC   library. The present sequence encodes a fusion protein of human RA12 with
CC   the protein product of a cDNA (full length, extended or partial) isolated
CC   from a library derived from lung tumour/cancer cells. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from the USPTO at
CC   seqdata.uspto.gov/sequence.html?DocId=20020197669
XX
SQ   Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                2.91e-15      Length:          186
Score:                    148.00        Matches:         30
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       7             Gaps:            0

US-09-684-215B-17 (1-30) x ACA12106 (1-186)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 3
ACA03292
ID   ACA03292 standard; cDNA; 186 BP.
XX
AC   ACA03292;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Lung cancer therapyand diagnosis associated cDNA #1767.
XX
KW   Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2002172952-A1.
XX
PD   21-NOV-2002.
XX
PF   10-JUL-2001; 2001US-00902941.
XX
PR   30-JUN-1999;   99US-00346492.
PR   15-OCT-1999;   99US-00419356.
PR   17-DEC-1999;   99US-00466867.
PR   30-DEC-1999;   99US-00476300.
PR   06-MAR-2000; 2000US-00519642.
PR   22-MAR-2000; 2000US-00533077.
PR   10-APR-2000; 2000US-00546259.
PR   27-APR-2000; 2000US-00560406.
PR   05-JUN-2000; 2000US-00589184.
PR   11-JUL-2000; 2000US-00614124.
PR   29-AUG-2000; 2000US-00651563.
PR   08-SEP-2000; 2000US-00658824.
PR   26-SEP-2000; 2000US-00671325.
PR   06-OCT-2000; 2000US-00677419.
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Henderson RA,  Wang T,  Watanabe Y,  Johnson JC,  Retter MW;
PI   Durham M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
XX
DR   WPI; 2003-328427/31.
XX
PT   New polynucleotide, useful for preparing a composition for treating or
PT   inhibiting development of cancer, e.g. lung cancer.
XX
PS   Example 10; SEQID NO 1879; 82pp; English.
XX
CC   The invention describes an isolated polynucleotide comprising one of 32
CC   sequences, complement or degenerate variants of them. The polynucleotide
CC   is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC   for treating or inhibiting development of cancer, e.g. lung cancer. This
CC   sequence represents a polynucleotide associated with the compositions and
CC   methods for the therapy and diagnosis of lung cancer
XX
SQ   Sequence 186 BP; 39 A; 61 C; 47 G; 39 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.91e-15      Length:          186
Score:                  148.00        Matches:         30
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     7             Gaps:            0

US-09-684-215B-17 (1-30) x ACA03292 (1-186)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 4
AAL40769
ID   AAL40769 standard; DNA; 396 BP.
XX
AC   AAL40769;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12 protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .396
FT                   /*tag=  a
FT                   /product= "Ra12 protein"
FT                   /note= "No start or stop codon"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22138.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Claim 1; Fig 2; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12 protein
XX
SQ   Sequence 396 BP; 68 A; 132 C; 134 G; 62 T; 0 U; 0 Other;

Alignment Scores:
```

```
Pred. No.:              7.45e-15        Length:         396
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-17 (1-30) x AAL40769 (1-396)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 60

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         61 GGGCAGGCGATGGCGATCGCGGGCCAGATC 90

RESULT 5
ABS71801
ID   ABS71801 standard; DNA; 399 BP.
XX
AC   ABS71801;
XX
DT   29-AUG-2003  (revised)
DT   02-DEC-2002  (first entry)
XX
DE   Human DNA encoding a mammaglobin/RA12 fusion protein.
XX
KW   Human; epitope; mammaglobin; breast cancer; cytostatic; T cell; CD4+;
KW   CD8+; antigen; RA12; ds; gene.
XX
OS   Homo sapiens.
OS   Mycobacterium tuberculosis.
OS   Chimeric.
XX
PN   WO200253017-A2.
XX
PD   11-JUL-2002.
XX
PF   08-JAN-2002; 2002WO-US003057.
XX
PR   08-JAN-2001; 2001US-00757417.
PR   08-NOV-2001; 2001US-00008045.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Fling SP,  Foy TM,  Clapper JD,  Wang A,  Johnson JC,  Mcneill PD;
PI   Sutherland RA;
XX
DR   WPI; 2002-706844/76.
DR   P-PSDB; ABG94684.
XX
PT   Novel polypeptides comprise one or more human mammoglobin epitopes and
PT   polynucleotides encoding the polypeptides, useful for preventing and
PT   treating breast cancers.
XX
PS   Claim 20; Page 115; 121pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) comprising 7-30
CC   consecutive amino acid residues of human mammaglobin, where one or more
CC   mammaglobin-specific T cells specifically reacts with (I). Also included
CC   are (1) a composition comprising (I), in combination with a
CC   physiologically acceptable carrier or immunostimulant; (2) a diagnostic
CC   kit, comprising (I) and a detection reagent comprising a reporter group;
CC   (3) removing tumour cells from a biological sample, by contacting a
CC   biological sample with T cells that specifically react with (I), under
CC   conditions and for a time sufficient to permit the removal of cells
CC   expressing mammaglobin or a peptide epitope from the sample; (4) an
CC   isolated T cell population (II), comprising T cells prepared using (I);
CC   and (5) a composition (C2) comprising a polynucleotide comprising a
CC   sequence encoding human mammaglobin (or variant or tagged with an
CC   affinity tag), or a polypeptide comprising a human mammaglobin (or
CC   variant or tagged with an affinity tag), in combination with an
CC   immunostimulant. The peptide is useful for inhibiting the development of
CC   breast cancer in a patient, by incubating CD4^+ and/or CD8+ T cells
CC   isolated from a patient with a mammoglobin epitope, such that T cells
CC   proliferate, administering the proliferated T cells to the patient,
CC   optionally cloning at least one proliferated T cell and administering at
CC   least one cloned cell, and thus inhibiting the development of breast
CC   cancer in the patient. The compositions and methods are useful for
CC   inhibiting the development of breast cancer in a patient. The present
CC   sequence encodes a human mammaglobin/RA12 fusion protein. (Updated on 29-
CC   AUG-2003 to standardise OS field)
XX
SQ   Sequence 399 BP; 109 A; 96 C; 92 G; 102 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.52e-15        Length:         399
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x ABS71801 (1-399)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 6
AAD60769
ID   AAD60769 standard; DNA; 399 BP.
XX
AC   AAD60769;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Ra12 (s)-human MammFL chimeric DNA construct.
XX
KW   Cytostatic; epitope; mammaglobin; mgb; therapy; breast cancer; diagnosis;
KW   vaccine; human; chimeric; fusion protein; ds.
XX
OS   Chimeric- Homo sapiens.
OS   Chimeric- Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .399
FT                   /*tag=  a
FT                   /product= "Ra12 (s)-human MammFL fusion protein"
XX
PN   US2002082216-A1.
XX
PD   27-JUN-2002.
XX
PF   08-JAN-2001; 2001US-00757417.
XX
PR   26-MAY-2000; 2000US-00580376.
XX
PA   (FANG/) FANGER G R.
PA   (FOYT/) FOY T M.
PA   (HOUG/) HOUGHTON R L.
PA   (REED/) REED S G.
XX
PI   Fanger GR,  Foy TM,  Houghton RL,  Reed SG;
XX
DR   WPI; 2003-776615/73.
DR   P-PSDB; ABW00067.
XX
PT   New polypeptide epitopes of human mammaglobin useful in inhibiting
PT   development of breast cancer and in breast cancer diagnosis and
PT   monitoring, and to produce antibodies also useful in breast cancer
```

```
PT   therapy, diagnosis and monitoring.
XX
PS   Example 8; Page 29; 66pp; English.
XX
CC   The present invention provides polypeptide epitopes of human mammaglobin
CC   (mgb) useful for the therapy, diagnosis and monitoring of breast cancer.
CC   The invention is related to specific epitopes of mammaglobin, to
CC   antibodies and immune cells that recognise such epitopes and to methods
CC   for detecting mammaglobin in patient serum. These peptides, antibodies
CC   and cells may be useful in vaccines and pharmaceutical compositions for
CC   prevention and treatment of breast cancer. The invention is also useful
CC   to detect and /or monitor the progression of breast cancer. The present
CC   sequence is Ra12 (s)-human MammFL chimeric DNA construct used in the
CC   expression of recombinant Ra12(s) MammFL, a fusion protein consisting of
CC   full length human mammaglobin with short Ra12. This protein is used in
CC   the exemplification of the invention
XX
SQ   Sequence 399 BP; 109 A; 96 C; 92 G; 102 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                7.52e-15       Length:          399
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       9              Gaps:            0

US-09-684-215B-17 (1-30) x AAD60769 (1-399)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 7
AAT91403
ID   AAT91403 standard; DNA; 447 BP.
XX
AC   AAT91403;
XX
DT   12-JAN-1998  (first entry)
XX
DE   Mycobacterium tuberculosis antigen TbRa12 encoding DNA.
XX
KW   Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW   skin testing; M.tuberculosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             11. .409
FT                   /*tag=  a
FT                   /product= "Antigen_TbRa12"
XX
PN   WO9709429-A2.
XX
PD   13-MAR-1997.
XX
PF   30-AUG-1996;   96WO-US014675.
XX
PR   01-SEP-1995;   95US-00523435.
PR   22-SEP-1995;   95US-00532136.
PR   22-MAR-1996;   96US-00620280.
PR   05-JUN-1996;   96US-00658800.
PR   12-JUL-1996;   96US-00680573.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Skeiky YAW,  Dillon DC,  Campos-Neto A,  Houghton R;
PI   Vedvick TH,  Twardzik DR;
XX
DR   WPI; 1997-192904/17.
DR   P-PSDB; AAW32354.
XX
PT   New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -
PT   useful for diagnosis of M. tuberculosis infection.
XX
PS   Claim 3; Page 49-50; 190pp; English.
XX
CC   A new immunogenic polypeptide has been developed comprising an
CC   immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC   variant differing only in conservative substitutions and/or
CC   modifications). The present sequence encodes a specifically claimed
CC   M.tuberculosis antigen, TbRa12. The immunogenic polypeptide can be used
CC   to diagnose M.tuberculosis infection by forming complexes with specific
CC   antibodies in the sample. Fragments of DNA encoding the immunogenic
CC   polypeptide can be used as diagnostic primers or probes and agents that
CC   bind to the antigen, especially monoclonal antibodies or equivalent
CC   polyclonal antibodies, are also used for diagnosis
XX
SQ   Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                8.66e-15       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       2              Gaps:            0

US-09-684-215B-17 (1-30) x AAT91403 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 8
AAT91466
ID   AAT91466 standard; DNA; 447 BP.
XX
AC   AAT91466;
XX
DT   07-JAN-1998  (first entry)
XX
DE   Mycobacterium tuberculosis antigen TbRa12 encoding DNA.
XX
KW   Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW   skin testing; M.tuberculosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             11. .409
FT                   /*tag=  a
FT                   /product= "Antigen_TbRa12"
XX
PN   WO9709428-A2.
XX
PD   13-MAR-1997.
XX
PF   30-AUG-1996;   96WO-US014674.
XX
PR   01-SEP-1995;   95US-00523436.
PR   22-SEP-1995;   95US-00533634.
PR   22-MAR-1996;   96US-00620874.
PR   05-JUN-1996;   96US-00659683.
PR   12-JUL-1996;   96US-00680574.
XX
PA   (CORI-) CORIXA CORP.
```

```
XX
PI   Reed SG,  Skeiky YA,  Dillon DC,  Campos-Neto A,  Houghton R;
PI   Vedvick TH,  Twardzik DR;
XX
DR   WPI; 1997-192903/17.
DR   P-PSDB; AAW32422.
XX
PT   New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
PT   useful in vaccines for prevention or treatment of tuberculosis, also for
PT   diagnosis.
XX
PS   Claim 3; Page 50; 168pp; English.
XX
CC   A new immunogenic polypeptide has been developed comprising an
CC   immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC   variant differing only in conservative substitutions and/or
CC   modifications). The present sequence encodes a specifically claimed
CC   M.tuberculosis antigen, TbRa12. The immunogenic protein, and fusion
CC   proteins containing one or more of the proteins or one of the proteins
CC   plus ESAT-6, are useful in vaccines, preferably when formulated with a
CC   non-specific adjuvant, to induce an immune response against
CC   M.tuberculosis (for treatment or prevention)
XX
SQ   Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              8.66e-15        Length:         447
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-09-684-215B-17 (1-30) x AAT91466 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 9
AAV44342
ID   AAV44342 standard; DNA; 447 BP.
XX
AC   AAV44342;
XX
DT   17-OCT-2003  (revised)
DT   09-NOV-1998  (first entry)
XX
DE   Mycobacterium tuberculosis antigen TbRa12 DNA.
XX
KW   Tuberculosis; infection; diagnosis; antigen; TbRa12; ss.
XX
OS   Mycobacterium tuberculosis; strain H37Ra.
XX
FH   Key             Location/Qualifiers
FT   CDS             11. .406
FT                   /*tag=  a
XX
PN   WO9816645-A2.
XX
PD   23-APR-1998.
XX
PF   07-OCT-1997;   97WO-US018214.
XX
PR   11-OCT-1996;   96US-00729622.
PR   13-MAR-1997;   97US-00818111.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Skeiky YAW,  Dillon DC,  Campos-Neto A,  Houghton R;
PI   Vedvick TS,  Twardzik DR,  Lodes MJ;
XX
DR   WPI; 1998-251292/22.
DR   P-PSDB; AAW64294.
XX
PT   New isolated Mycobacterium tuberculosis polypeptides and DNA - used to
PT   develop products for the detection of M. tuberculosis infection and
PT   diagnosis of tuberculosis.
XX
PS   Claim 3; Page 61; 250pp; English.
XX
CC   This DNA sequence codes for Mycobacterium tuberculosis soluble antigen
CC   TbRa12 (see AAW64294). It was isolated from a M. tuberculosis strain
CC   H37Ra expression library with rabbit anti-sera raised against M.
CC   tuberculosis supernatant. The invention relates to compositions and
CC   methods for diagnosing tuberculosis. It provides polypeptides (see
CC   AAW64291-W64379) comprising an antigenic portion of a soluble M.
CC   tuberculosis antigen, or an immunogenic portion of an M. tuberculosis
CC   antigen, as well as DNA sequences encoding such polypeptides, recombinant
CC   expression vectors and transformed or transfected host cells. Also
CC   claimed are methods and diagnostic kits for detecting M. tuberculosis
CC   infection in a patient using these polypeptides, antibodies or
CC   oligonucleotide probes and primers, for the diagnosis of tuberculosis.
CC   (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              8.66e-15        Length:         447
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     2               Gaps:           0

US-09-684-215B-17 (1-30) x AAV44342 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 10
AAV64450
ID   AAV64450 standard; DNA; 447 BP.
XX
AC   AAV64450;
XX
DT   27-JAN-1999  (first entry)
XX
DE   M. tuberculosis immunogenic polypeptide TbRa12 DNA.
XX
KW   Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KW   vaccine; pharmaceutical; infection; diagnosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9816646-A2.
XX
PD   23-APR-1998.
XX
PF   07-OCT-1997;   97WO-US018293.
XX
PR   11-OCT-1996;   96US-00730510.
PR   13-MAR-1997;   97US-00818112.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Reed SG,  Skeiky YAW,  Dillon DC,  Campos-Neto A,  Houghton R;
```

PI Vedvick TS, Twardzik DR, Lodes MJ;
XX
DR WPI; 1998-261042/23.
DR P-PSDB; AAW81657.
XX
PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
PT develop products for the detection of M. tuberculosis infection and for
PT diagnosis, treatment and prevention of tuberculosis.
XX
PS Claim 3; Page 62-63; 230pp; English.
XX
CC This sequence encodes an immunogenic portion of a soluble Mycobacterium
CC tuberculosis (MT) antigen which can be used in a method for inducing
CC protective immunity against tuberculosis (TB). This sequence can be
CC formulated into vaccines and/or pharmaceutical compositions for
CC immunising against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	8.66e-15	Length:	447
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-684-215B-17 (1-30) x AAV64450 (1-447)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100
```

RESULT 11
AAZ19040
ID AAZ19040 standard; DNA; 447 BP.
XX
AC AAZ19040;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis recombinant antigen DNA encoding TbRa12.
XX
KW Antigen; diagnosis; detection; infection; antibody; immunisation;
KW vaccine; immunity; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9942118-A2.
XX
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US003265.
XX
PR 18-FEB-1998; 98US-00024753.
PR 05-MAY-1998; 98US-00072596.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
DR WPI; 1999-527416/44.
DR P-PSDB; AAY38959.
XX
PT New polypeptide comprising antigenic portions of M. tuberculosis.
XX
PS Claim 3; Page 103; 323pp; English.
XX
CC This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against M.
CC tuberculosis infection. The new detection methods are needed as current
CC vaccination strategies do not provide 100% immunity
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	8.66e-15	Length:	447
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-684-215B-17 (1-30) x AAZ19040 (1-447)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100
```

RESULT 12
AAZ19252
ID AAZ19252 standard; DNA; 447 BP.
XX
AC AAZ19252;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis antigen TbRa12 DNA sequence.
XX
KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9942076-A2.
XX
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99WO-US003268.
XX
PR 18-FEB-1998; 98US-00025197.
PR 05-MAY-1998; 98US-00072967.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX
DR WPI; 1999-527409/44.
DR P-PSDB; AAY39096.
XX
PT New antigens from Mycobacterium tuberculosis useful in diagnostic skin
PT tests and protective or therapeutic vaccines or compositions.
XX
PS Claim 3; Page 73; 299pp; English.
XX
CC The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are
CC vaccines and fusion protein containing M. tuberculosis Ag's. M.
CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other
CC polypeptides fragments, can be used in pharmaceutical compositions or
CC vaccines to generate a protective or therapeutic immune response to M.

```
CC   tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.
CC   Ag can induce proliferation of, or cytokine secretion by, T, B or natural
CC   killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249
CC   to AAZ19460 and AAY39083 to AAY39225 are used in the exemplification of
CC   the present invention
XX
SQ   Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                8.66e-15       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       2              Gaps:            0

US-09-684-215B-17 (1-30) x AAZ19252 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 13
AAS03780
ID   AAS03780 standard; DNA; 447 BP.
XX
AC   AAS03780;
XX
DT   29-AUG-2001  (first entry)
XX
DE   M. tuberculosis DNA encoding a partial antigen TbRa12.
XX
KW   TbRa12; antigen; vaccine; tuberculosis; AIDS;
KW   acquired immunodeficiency disease; ss.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             11. .409
FT                   /*tag=  a
FT                   /product= "TbRa12"
FT                   /partial
FT                   /note= "No start codon"
XX
PN   WO200124820-A1.
XX
PD   12-APR-2001.
XX
PF   10-OCT-2000; 2000WO-US028095.
XX
PR   07-OCT-1999;   99US-0158338P.
PR   07-OCT-1999;   99US-0158425P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Reed S,  Houghton RL,  Mcneill PD,  Dillon DC,  Lodes ML;
XX
DR   WPI; 2001-290576/30.
DR   P-PSDB; AAU01889.
XX
PT   Vaccinating against Mycobacteria infections in mammals using fusion
PT   proteins comprising combinations of heterologous antigens.
XX
PS   Example 2; Page 162; 168pp; English.
XX
CC   The sequence encodes Mycobacterium tuberculosis TbRa12, an M.
CC   tuberculosis antigen. Compositions comprising at least 2 heterologous
CC   antigens, as a fusion protein, and vectors expressing the fusion proteins
CC   are used as vaccines to prophylactically immunise mammals (especially
CC   humans) against infection by Mycobacteria. The compositions contain at
CC   least 2 heterologous antigens that increase the serological sensitivity
CC   of individuals infected with tuberculosis, a disease frequently affecting
CC   patients with acquired immunodeficiency disease, AIDS
XX
SQ   Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:                8.66e-15       Length:          447
Score:                    148.00         Matches:         30
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              100.00%        Indels:          0
DB:                       5              Gaps:            0

US-09-684-215B-17 (1-30) x AAS03780 (1-447)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100

RESULT 14
AAD47080
ID   AAD47080 standard; DNA; 447 BP.
XX
AC   AAD47080;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Mycobacterium tuberculosis Ra12 antigen encoding DNA.
XX
KW   Vaccine; immunity; diagnostic agent; gene therapy; Ra12 antigen; gene;
KW   ds.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key             Location/Qualifiers
FT   CDS             11. .409
FT                   /*tag=  a
FT                   /product= "Ra12 antigenic protein"
FT                   /note= "No start codon"
FT                   /partial
XX
PN   WO200272792-A2.
XX
PD   19-SEP-2002.
XX
PF   13-MAR-2002; 2002WO-US008223.
XX
PR   13-MAR-2001; 2001US-0275837P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Brannon M,  Guderian J;
XX
DR   WPI; 2002-759844/82.
DR   P-PSDB; AAE29705.
XX
PT   New recombinant nucleic acid molecule comprising a Leishmania TSA, LeIF,
PT   M15 or 6H polynucleotide, useful as vaccine to elicit protective immunity
PT   against pathogenic microorganisms e.g. Leishmania and Mycobacterium
PT   tuberculosis.
XX
PS   Disclosure; Page 83; 155pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide. The recombinant nucleic acid comprises a heterologous
CC   polynucleotide sequence encoding an antigen or an antigenic fragment from
CC   Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
```

CC polypeptide or its fragment. The Leishmania polynucleotide is selected
CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
CC are used in methods for eliciting immune response in mammals. They are
CC useful as vaccines to elicit protective immunity against pathogenic
CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
CC polypeptides are used for enhancing the expression of polynucleotides, as
CC in vivo diagnostic agents and for raising antibodies in a non-human
CC animal. The invention is used in gene therapy. The present sequence is M.
CC tuberculosis Ra12 (C-terminus of MTB32A; RA35FL) antigen encoding DNA
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	8.66e-15	Length:	447
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x AAD47080 (1-447)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100
```

RESULT 15
AAD28339
ID AAD28339 standard; cDNA; 447 BP.
XX
AC AAD28339;
XX
DT 22-APR-2002 (first entry)
XX
DE Mycobacterium species MTB32A C-terminal peptide encoding cDNA, MTBRa12.
XX
KW Fusion protein; antigen; serological sensitivity; immune response;
KW tuberculosis; infection; vaccine; MTB32A; Ra32FL; MTBRa12; ss.
XX
OS Mycobacterium sp.
XX
FH Key Location/Qualifiers
FT CDS 11. .409
FT /*tag= a
FT /product= "Ra35 protein fragment"
FT /note= "CDS does not include start codon"
FT /partial
XX
PN WO200198460-A2.
XX
PD 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-US019959.
XX
PR 20-JUN-2000; 2000US-00597796.
PR 01-FEB-2001; 2001US-0265737P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Reed S, Alderson M;
XX
DR WPI; 2002-147798/19.
DR P-PSDB; AAE17569.
XX
PT Composition comprising MTB39 antigen and MTB32A antigen from
PT Mycobacterium species, useful for eliciting immune response in a subject.
XX
PS Disclosure; Page 99; 136pp; English.
XX

CC The present invention relates to fusion proteins containing at least two
CC Mycobacterium species antigens, nucleotides encoding them and
CC compositions comprising such fusion proteins. The present invention
CC particularly relates to nucleic acids encoding fusion proteins that
CC include two or more individual M. tuberculosis antigens which increase
CC the serological sensitivity of sera from individuals infected with
CC tuberculosis and methods for their use in diagnosis, prevention and
CC treatment of tuberculosis infection. Sequences of the invention are
CC useful for eliciting an immune response in a mammal, e.g., human,
CC immunised with BCG. They are useful in the diagnosis, treatment and
CC prevention of Mycobacterium infection. The fusion proteins and the
CC polynucleotides are useful as diagnostic tools in patients infected with
CC Mycobacterium, in vitro and in vivo assays for detecting humoral
CC antibodies or cell-mediated immunity against M. tuberculosis, for the
CC diagnosis of an infection or monitoring of disease progression, as
CC immunogens to generate or elicit a protective immune response in a
CC patient and for raising anti-M. tuberculosis antibodies in a non-human
CC animal. Sequences of the invention are also used as vaccines. MTB32A
CC fusion proteins of the invention are useful as in vivo diagnostic agents
CC for intradermal skin test. The present sequence is Mycobacterium species
CC MTB32A (Ra32FL) C-terminal protein fragment (residues 244-355) encoding
CC cDNA, MTBRa12
XX
SQ Sequence 447 BP; 79 A; 146 C; 149 G; 72 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	8.66e-15	Length:	447
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x AAD28339 (1-447)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||
Db         11 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 70

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         71 GGGCAGGCGATGGCGATCGCGGGCCAGATC 100
```

RESULT 16
AAL40772
ID AAL40772 standard; DNA; 672 BP.
XX
AC AAL40772;
XX
DT 06-AUG-2003 (revised)
DT 03-OCT-2002 (first entry)
XX
DE Nucleotide sequence encoding Ra12-mammaglobin fusion protein.
XX
KW Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW immunogen; cytokine; gene; ds.
XX
OS Mammalia.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 4. .666
FT /*tag= a
FT /product= "Ra12-mammaglobin fusion protein"
XX
PN WO200125401-A2.
XX
PD 12-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US027652.
XX

PR 07-OCT-1999; 99US-0158585P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky Y, Guderian J;
XX
DR WPI; 2001-266299/27.
DR P-PSDB; AAO22141.
XX
PT Recombinant nucleic acid molecule for producing high yield expression of
PT desired fusion polypeptides, encodes fusion polypeptide comprising
PT Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS Disclosure; Fig 5; 39pp; English.
XX
CC The invention relates to a recombinant nucleic acid molecule encoding a
CC fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC kDa C-terminal fragment of serine protease antigen MTB32A of
CC Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC The recombinant fusion nucleic acids and polypeptides are useful for
CC providing stable and high yield expression of fusion polypeptides of both
CC eukaryotic and prokaryotic origin and to encode a protein product for use
CC as an antigen for detecting serum antibodies. The presence of serum
CC antibodies to M. tuberculosis antigens in an individual indicates that
CC the individual is infected with it. The fusion polypeptides are useful as
CC sources of proteins for monitoring binding of serum antibodies to fusion
CC proteins and as an immunogen to induce and/or enhance immune responses.
CC The coding sequences can be ligated with a coding sequence of another
CC molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC represents the DNA encoding the Ra12-mammaglobin fusion protein. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 672 BP; 160 A; 185 C; 187 G; 140 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.44e-14	Length:	672
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x AAL40772 (1-672)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114
```

RESULT 17
AAH93896
ID AAH93896 standard; cDNA; 675 BP.
XX
AC AAH93896;
XX
DT 04-OCT-2001 (first entry)
XX
DE Ra12-P510S-C construct cDNA sequence.
XX
KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytostatic; gene therapy; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US001574.
XX
PR 14-JAN-2000; 2000US-00483672.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX
DR WPI; 2001-425873/45.
XX
PT New polynucleotide encoding a prostate-specific protein, for diagnosing,
PT monitoring and treating prostate cancer in a patient and for use in
PT vaccines.
XX
PS Claim 8; Page 492-493; 543pp; English.
XX
CC The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated T cells
CC prepared using (I) or (II) are used treat cancer in a patient. (I) and
CC the antibodies are also used in the detection of cancer in a patient. The
CC cancer that is diagnosed or treated is particularly prostate cancer. (I)
CC and (II) can be used in vaccines. The antibodies or (I) can be used for
CC monitoring the progression of cancer in a patient. (I) and (II) can also
CC be used to improve diagnostic and therapeutic methods for prostate
CC cancer. They can indicate the level of metastasis as well as the prostate
CC volume. AAH93357 to AAH93944 and AAM01115 to AAM01318 represent
CC polynucleotide and amino acid sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.45e-14	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x AAH93896 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 18
AAS64132
ID AAS64132 standard; cDNA; 675 BP.
XX
AC AAS64132;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human /M. tuberculosis Ra12 fusion protein RA12-P510S-C cDNA.
XX
KW Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS Homo sapiens.
OS Microbacterium; tuberculosis.
OS Synthetic.
OS Chimeric.
XX
PN WO200173032-A2.
XX
PD 04-OCT-2001.
XX

PF 27-MAR-2001; 2001WO-US009919.
XX
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR WPI; 2001-639232/73.
DR P-PSDB; AAU69899.
XX
PT New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer.
XX
PS Example 17; Page 532; 579pp; English.
XX
CC The invention relates to isolated prostate-specific polynucleotides,
CC polypeptides, fusion proteins of the polypeptides, antibodies raised
CC against the polypeptides (or antigenic epitopes derived from them) and
CC antigen-presenting cells expressing the polypeptides. The antibodies are
CC useful for detecting the presence of cancer, especially prostate cancer.
CC The polypeptides, polynucleotides and the antigen-presenting cells are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein, and for inhibiting the development of cancer especially prostate
CC cancer. Compositions comprising the polynucleotide and/or polypeptide are
CC useful for stimulating an immune response, and for treating cancer. The
CC oligonucleotide is useful for detecting cancer. The present sequence is a
CC prostate specific polynucleotide of the invention
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.45e-14	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-684-215B-17 (1-30) x AAS64132 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 19
ACA59940
ID ACA59940 standard; cDNA; 675 BP.
XX
AC ACA59940;
XX
DT 10-JUN-2003 (first entry)
XX
DE Prostate cancer therapy associated cDNA #647.
XX
KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;
KW immunogen; cancer; prostate specific antigen; PSA;
KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;
KW PSMA; gene; ss.
XX
OS Homo sapiens.
XX
PN US2002192763-A1.
XX
PD 19-DEC-2002.
XX
PF 29-JUN-2001; 2001US-00895793.
XX
PR 04-OCT-1999; 99US-0157455P.
PR 04-OCT-2000; 2000US-00679272.
PR 28-MAR-2001; 2001US-00822827.
XX
PA (XUJJ/) XU J.
PA (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (HARL/) HARLOCKER S L.
PA (JIAN/) JIANG Y.
PA (KALO/) KALOS M D.
PA (FANG/) FANGER G R.
PA (RETT/) RETTER M W.
PA (STOL/) STOLK J A.
PA (DAYC/) DAY C H.
PA (VEDV/) VEDVICK T S.
PA (CART/) CARTER D.
PA (LISX/) LI S X.
PA (WANG/) WANG A.
PA (SKEI/) SKEIKY Y A W.
PA (HEPL/) HEPLER W T.
PA (HEND/) HENDERSON R A.
PA (HURA/) HURAL J.
PA (MCNE/) MCNEILL P D.
PA (HOUG/) HOUGHTON R L.
PA (DBAS/) Y DE BASSOLS C V.
PA (FOYT/) FOY T M.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;
PI Mcneill PD, Houghton RL, Y De BassolsCV, Foy TM;
XX
DR WPI; 2001-245062/25.
XX
PT Prostate specific protein and its encoding polynucleotide, useful for the
PT treatment and diagnosis of prostate cancer.
XX
PS Example 17; SEQ ID NO 822; 85pp; English.
XX
CC The invention describes a fusion protein comprising at least one amino
CC acid sequence of immunogenic portions of any of the 3 sequences not
CC defined in the specification, or sequences having at least 70 or 90 %
CC sequence identity to any one of the 35 sequences defined in the USPTO web
CC site, which is encoded by any of the 4 nucleotide sequences not defined
CC in the specification. The fusion protein, composition and methods are
CC useful for diagnosing, preventing and/or treating cancer, particularly
CC prostate cancer. The proteins are useful as markers to indicate the
CC presence or absence of cancer. This sequence represents a prostate cancer
CC therapy associated cDNA. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=US20020192763
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.45e-14	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

```
US-09-684-215B-17 (1-30) x ACA59940 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 20
ABL95503
ID   ABL95503 standard; cDNA; 675 BP.
XX
AC   ABL95503;
XX
DT   29-AUG-2003  (revised)
DT   19-JUL-2002  (first entry)
XX
DE   Ra12- P510S-C construct cDNA sequence SEQ ID NO 822.
XX
KW   Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
KW   gene therapy; gene; ss.
XX
OS   Mycobacterium tuberculosis.
OS   Homo sapiens.
OS   Chimeric.
XX
PN   US2002022248-A1.
XX
PD   21-FEB-2002.
XX
PF   12-JAN-2001; 2001US-00759143.
XX
PR   25-FEB-1997;   97US-00806099.
PR   01-AUG-1997;   97US-00904804.
PR   10-FEB-1998;   98US-00020956.
PR   25-FEB-1998;   98US-00030607.
PR   14-JUL-1998;   98US-00115453.
PR   23-SEP-1998;   98US-00159812.
PR   15-JAN-1999;   99US-00232149.
PR   09-APR-1999;   99US-00288946.
PR   13-JUL-1999;   99US-00352616.
PR   12-NOV-1999;   99US-00439313.
PR   18-NOV-1999;   99US-00443686.
PR   14-JAN-2000; 2000US-00483672.
PR   27-MAR-2000; 2000US-00536857.
PR   09-MAY-2000; 2000US-00568100.
PR   12-MAY-2000; 2000US-00570737.
PR   13-JUN-2000; 2000US-00593793.
PR   27-JUN-2000; 2000US-00605783.
PR   10-AUG-2000; 2000US-00636215.
PR   29-AUG-2000; 2000US-00651236.
PR   06-SEP-2000; 2000US-00657279.
PR   02-OCT-2000; 2000US-00679426.
PR   10-OCT-2000; 2000US-00685166.
XX
PA   (XUJJ/) XU J.
PA   (DILL/) DILLON D C.
PA   (MITC/) MITCHAM J L.
PA   (HARL/) HARLOCKER S L.
PA   (JIAN/) JIANG Y.
PA   (KALO/) KALOS M D.
PA   (FANG/) FANGER G R.
PA   (RETT/) RETTER M W.
PA   (STOL/) STOLK J A.
PA   (DAYC/) DAY C H.
PA   (VEDV/) VEDVICK T S.
PA   (CART/) CARTER D.
PA   (LISX/) LI S X.
PA   (WANG/) WANG A.
PA   (SKEI/) SKEIKY Y A W.
PA   (HEPL/) HEPLER W T.
PA   (HEND/) HENDERSON R A.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI   Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI   Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
XX
DR   WPI; 2002-255649/30.
XX
PT   New prostate-specific polynucleotides for diagnosing and treating
PT   diseases, in particular prostate cancer, and as markers for the
PT   progression of cancer.
XX
PS   Example 17; SEQ ID NO 822; 87pp; English.
XX
CC   The present invention provides prostate-specific coding sequences and
CC   their encoded proteins. These can be used in the diagnosis and treatment
CC   of cancers, particularly prostate cancer. The present sequence is a cDNA
CC   described in the invention. (Updated on 29-AUG-2003 to standardise OS
CC   field)
XX
SQ   Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.45e-14        Length:          675
Score:                    148.00          Matches:         30
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-17 (1-30) x ABL95503 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 21
ACC95667
ID   ACC95667 standard; cDNA; 675 BP.
XX
AC   ACC95667;
XX
DT   28-AUG-2003  (first entry)
XX
DE   Prostate tumour specific cDNA sequence SEQ ID 822.
XX
KW   Cytostatic; gene therapy; prostate-specific protein; PSP; human;
KW   immune response; prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200289747-A2.
XX
PD   14-NOV-2002.
XX
PF   09-MAY-2002; 2002WO-US014753.
XX
PR   09-MAY-2001; 2001US-00852911.
PR   29-JUN-2001; 2001US-00895814.
PR   10-DEC-2001; 2001US-00012896.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA;
PI   Kalos MD, Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS;
PI   Carter D, Li SX, Wang A, Skeiky YAW, Hepler WT, Hural J;
PI   Mcneill PD, Houghton RL, Vinals Y De BassolsC, Foy TM, Watanabe Y;
PI   Deng T;
```

XX
DR WPI; 2003-167130/16.
XX
PT New prostate-specific proteins and genes, useful in gene therapy,
PT particularly for stimulating an immune response in a patient, or treating
PT prostate cancer in a patient, as well as for diagnosing prostate cancer
PT in a patient.
XX
PS Example 17; Page 601-602; 691pp; English.
XX
CC The present invention relates to novel prostate-specific proteins (PSP)
CC and their coding sequences. The PSPs and their coding sequences are
CC useful for stimulating an immune response in a patient, or for treating
CC prostate cancer in a patient and for determining, detecting or diagnosing
CC the presence of a cancer in a patient. The present sequence was used to
CC illustrate the invention
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.45e-14	Length:	675
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-09-684-215B-17 (1-30) x ACC95667 (1-675)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 22
ADB14272
ID ADB14272 standard; cDNA; 675 BP.
XX
AC ADB14272;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human prostate specific protein P510S-RA12 construct C cDNA.
XX
KW Human; ss; prostate specific cDNA; cytostatic; immunostimulant;
KW gene therapy; cell therapy; vaccine; T-cell epitope;
KW class I major histocompatibility complex allele; MHC; prostate cancer;
KW tumour; antigen presenting cell; gene; fusion protein.
XX
OS Chimeric.
OS Homo sapiens.
OS Mycobacterium tuberculosis.
XX
PN US2003185830-A1.
XX
PD 02-OCT-2003.
XX
PF 12-NOV-2002; 2002US-00294025.
XX
PR 25-FEB-1997; 97US-00806099.
PR 01-AUG-1997; 97US-00904804.
PR 09-FEB-1998; 98US-00020956.
PR 25-FEB-1998; 98US-00030607.
PR 14-JUL-1998; 98US-00115453.
PR 23-SEP-1998; 98US-00159812.
PR 15-JAN-1999; 99US-00232149.
PR 09-APR-1999; 99US-00288946.
PR 13-JUL-1999; 99US-00352616.
PR 12-NOV-1999; 99US-00439313.
PR 18-NOV-1999; 99US-00443686.
PR 14-JAN-2000; 2000US-00483672.
PR 27-MAR-2000; 2000US-00536857.
PR 09-MAY-2000; 2000US-00568100.
PR 12-MAY-2000; 2000US-00570737.
PR 13-JUN-2000; 2000US-00593793.
PR 27-JUN-2000; 2000US-00605783.
PR 09-AUG-2000; 2000US-00636215.
PR 29-AUG-2000; 2000US-00651236.
PR 06-SEP-2000; 2000US-00657279.
PR 02-OCT-2000; 2000US-00679426.
PR 10-OCT-2000; 2000US-00685166.
PR 09-NOV-2000; 2000US-00709729.
PR 12-JAN-2001; 2001US-00759143.
PR 09-FEB-2001; 2001US-00780669.
PR 09-MAY-2001; 2001US-00852911.
PR 29-JUN-2001; 2001US-00895814.
PR 10-DEC-2001; 2001US-00012896.
PR 09-MAY-2002; 2002US-00144678.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Stolk JA, Kalos MD;
XX
DR WPI; 2003-756193/71.
DR P-PSDB; ADB14275.
XX
PT New isolated polypeptide for use in a vaccine for stimulating an immune
PT response, or for treating or diagnosis cancer, preferably prostate
PT cancer.
XX
PS Example 17; Page; 101pp; English.
XX
CC The invention relates to an isolated polypeptide comprising no more than
CC 11-542 amino acids of ADB13563 comprising a sequence ADB14487. The
CC peptides comprise a fragment ADB13563 of that contain naturally processed
CC T-cell epitopes for 3 class I major histocompatibility complex (MHC)
CC alleles. ADB13563 is a polypeptide encoded by a human prostate specific
CC cDNA, one of 648 disclosed as new. Also included are nucleic acids
CC encoding the proteins and peptides, expression vectors, a host cell
CC transformed with the vector, an isolated antibody (or antigen binding
CC fragment) that specifically binds to the protein or peptide, detecting
CC the presence of a cancer in a patient (comprising contacting a patient
CC sample with a binding agent that binds to the peptides or a polypeptide
CC appearing as ADB13558, detecting the amount of polypeptide that binds to
CC the agent and comparing the amount of polypeptide to a predetermined cut-
CC off value to determine the presence of cancer), a fusion protein
CC comprising the peptides or proteins, stimulating or expanding T cells
CC specific for a tumour protein comprising contacting T cells with the
CC peptides or the isolated T cell population, treating prostate cancer in a
CC patient comprising administering a composition comprising the peptides,
CC nucleic acids, antibodies or compounds, determining the presence of a
CC cancer in a patient and treating prostate cancer in a patient comprising
CC incubating cluster of differentiation (CD4+) and/or CD8+ T cells isolated
CC from a patient with the peptides or antigen presenting cells that express
CC (the peptides so that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The peptides (or an oligonucleotide
CC that hybridises to nucleic acid encoding them), is used to detect the
CC presence of cancer in a patient. The peptides, nucleic acids encoding, or
CC antigen-presenting cells expressing the nucleic acid, are used to
CC stimulate or expand T cells specific for a tumour protein. The peptides,
CC nucleic acids, antibodies, fusion proteins, T cell populations or antigen
CC presenting cells are used to stimulate an immune response or treat
CC prostate cancer in a patient. The present sequence is a cDNA encoding a
CC fusion protein comprising a prostate specific protein. Note: Except where
CC otherwise indicated, the sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030185830.
XX
SQ Sequence 675 BP; 162 A; 197 C; 190 G; 126 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.45e-14	Length:	675
Score:	148.00	Matches:	30

```
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     9              Gaps:          0

US-09-684-215B-17 (1-30) x ADB14272 (1-675)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 23
AAZ20206
ID   AAZ20206 standard; DNA; 702 BP.
XX
AC   AAZ20206;
XX
DT   17-JAN-2000  (first entry)
XX
DE   Mycobacterium tuberculosis antigen fusion protein Mtb24 DNA.
XX
KW   Tuberculosis; antigen; fusion protein; Mtb24; Ra12; DPPD; diagnosis;
KW   therapy; vaccine; immunogen; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9951748-A2.
XX
PD   14-OCT-1999.
XX
PF   07-APR-1999;   99WO-US007717.
XX
PR   07-APR-1998;   98US-00056556.
PR   30-DEC-1998;   98US-00223040.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Alderson M,  Campos-Neto A;
XX
DR   WPI; 1999-601610/51.
DR   P-PSDB; AAY32071.
XX
PT   New fusion proteins useful for diagnosis, prevention and treatment of
PT   tuberculosis.
XX
PS   Example; Fig 13A-B; 83pp; English.
XX
CC   This DNA sequence includes a coding region for a recombinant
CC   Mycobacterium tuberculosis bi-antigen fusion protein (see AAY32071),
CC   termed Mtb24, composed of the antigens Ra12 and DPPD. The DNA is useful
CC   for the recombinant production of the fusion protein. Coding sequences
CC   for the antigens were modified by PCR in order to facilitate their fusion
CC   and subsequent expression of the fusion protein, and then ligated. The
CC   invention provides fusion proteins (see AAY32059-71) containing at least
CC   2 M. tuberculosis antigens. The new fusion proteins and polynucleotides
CC   encoding them are useful as vaccines for preventing tuberculosis
CC   (claimed), for diagnosis (via in vitro assays or intradermal skin tests
CC   for detection of anti-M. tuberculosis antibodies), monitoring of disease
CC   progression, and treatment of tuberculosis. They are more effective
CC   immunogens than mixtures of the individual protein components
XX
SQ   Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.52e-14       Length:        702
Score:                  148.00         Matches:       30
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     2              Gaps:          0

US-09-684-215B-17 (1-30) x AAZ20206 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114

RESULT 24
AAL40770
ID   AAL40770 standard; DNA; 702 BP.
XX
AC   AAL40770;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Nucleotide sequence encoding Ra12-DPPD fusion protein.
XX
KW   Ra12; serine protease antigen; MTB32A; Mycobacterium tuberculosis;
KW   vaccine; eukaryotic; prokaryotic origin; serum antibody; fusion protein;
KW   immunogen; cytokine; gene; ds.
XX
OS   Unidentified.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             4. .696
FT                   /*tag=  a
FT                   /product= "Ra12-DPPD fusion protein"
XX
PN   WO200125401-A2.
XX
PD   12-APR-2001.
XX
PF   06-OCT-2000; 2000WO-US027652.
XX
PR   07-OCT-1999;   99US-0158585P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky Y,  Guderian J;
XX
DR   WPI; 2001-266299/27.
DR   P-PSDB; AAO22139.
XX
PT   Recombinant nucleic acid molecule for producing high yield expression of
PT   desired fusion polypeptides, encodes fusion polypeptide comprising
PT   Mycobacterium tuberculosis coding sequence and heterologous polypeptide.
XX
PS   Example 1; Fig 3; 39pp; English.
XX
CC   The invention relates to a recombinant nucleic acid molecule encoding a
CC   fusion polypeptide, comprising a polynucleotide sequence of Ra12, a 14
CC   kDa C-terminal fragment of serine protease antigen MTB32A of
CC   Mycobacterium tuberculosis, and a heterologous polynucleotide sequence.
CC   The recombinant fusion nucleic acids and polypeptides are useful for
CC   providing stable and high yield expression of fusion polypeptides of both
CC   eukaryotic and prokaryotic origin and to encode a protein product for use
CC   as an antigen for detecting serum antibodies. The presence of serum
CC   antibodies to M. tuberculosis antigens in an individual indicates that
CC   the individual is infected with it. The fusion polypeptides are useful as
CC   sources of proteins for monitoring binding of serum antibodies to fusion
CC   proteins and as an immunogen to induce and/or enhance immune responses.
CC   The coding sequences can be ligated with a coding sequence of another
CC   molecule such as M. tuberculosis antigens, a cytokine or an adjuvant and
CC   can be used in vivo as a DNA vaccine. This polynucleotide sequence
CC   represents the DNA encoding the Ra12-DPPD fusion protein
XX
SQ   Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              1.52e-14        Length:         702
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     4               Gaps:           0

US-09-684-215B-17 (1-30) x AAL40770 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114

RESULT 25
ABK14140
ID   ABK14140 standard; DNA; 702 BP.
XX
AC   ABK14140;
XX
DT   29-AUG-2003  (revised)
DT   08-MAY-2002  (first entry)
XX
DE   DNA encoding antigenic fusion protein Ra12-DPPD (Mtb24).
XX
KW   Fusion protein; tuberculosis; Mycobacterium tuberculosis; gene; ds;
KW   tuberculostatic; immunogen; vaccine; Ra12-DPPD; Mtb24.
XX
OS   Mycobacterium tuberculosis.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .702
FT                   /*tag=  a
FT                   /product= "Mtb24 #1"
FT                   /partial
FT                   /transl_except= (pos:694. .696, aa:Xaa)
FT                   /note= "No start or stop codon. Xaa= In frame stop codon"
FT   CDS             2. .702
FT                   /*tag=  b
FT                   /product= "Mtb24 #2"
FT                   /partial
FT                   /note= "No start or stop codon"
FT                   /transl_except= (pos:263. .265, aa:Xaa) /transl_except=
FT                   (pos:353. .355, aa:Xaa) /transl_except= (pos:395. .397,
FT                   aa:Xaa) /transl_except= (pos:470. .472, aa:Xaa)
FT                   /transl_except= (pos:701. .702, aa:Ser)
FT                   /note= "This codon has an apparent 1 nucleotide deletion
FT                   which alters the reading frame. Xaa= In frame stop codon"
FT   CDS             3. .701
FT                   /*tag=  c
FT                   /product= "Mtb24 #3"
FT                   /partial
FT                   /note= "No start or stop codon"
FT                   /transl_except= (pos:1. .2, aa:Pro) /transl_except=
FT                   (pos:39. .41, aa:Xaa) /transl_except= (pos:321. .323,
FT                   aa:Xaa) /transl_except= (pos:339. .341, aa:Xaa)
FT                   /transl_except= (pos:450. .452, aa:Xaa) /transl_except=
FT                   (pos:621. .623, aa:Xaa)
FT                   /note= "No start or stop codon. Xaa= In frame stop codon"
XX
PN   US2002009459-A1.
XX
PD   24-JAN-2002.
XX
PF   07-APR-1999;   99US-00287849.
XX
PR   13-MAR-1997;   97US-00818112.
PR   01-OCT-1997;   97US-00942578.
PR   18-FEB-1998;   98US-00025197.
PR   07-APR-1998;   98US-00056556.
PR   30-DEC-1998;   98US-00223040.
XX
PA   (REED/) REED S G.
PA   (SKEI/) SKEIKY Y A.
PA   (DILL/) DILLON D C.
PA   (ALDE/) ALDERSON M.
PA   (CAMP/) CAMPOS-NETO A.
XX
PI   Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;
XX
DR   WPI; 2002-171134/22.
DR   P-PSDB; AAU74600, AAU76541, AAU76542.
XX
PT   New fusion proteins of Mycobacterium tuberculosis antigens, useful for
PT   diagnosing, treating or preventing M. tuberculosis infection,
PT   particularly as vaccine for treating or preventing tuberculosis.
XX
PS   Example; Fig 13; 62pp; English.
XX
CC   The invention relates to a purified polypeptide which induces an immune
CC   response of Mycobacterium tuberculosis. Polypeptides of the invention are
CC   useful for diagnosing, treating or preventing M. tuberculosis infection,
CC   particularly tuberculosis infection. In particular, the polypeptides are
CC   useful as a vaccine formulation with an adjuvant to afford long-term
CC   protection in animals against the development of tuberculosis. The
CC   protein coding sequence may be used to encode a protein product for use
CC   as an immunogen to induce and/or enhance an immune response to M.
CC   tuberculosis. This sequence represents DNA encoding an M. tuberculosis
CC   fusion protein of the invention. This polynucleotide encodes 3 different
CC   proteins, each in a different reading frame. (Updated on 29-AUG-2003 to
CC   standardise OS field)
XX
SQ   Sequence 702 BP; 127 A; 229 C; 225 G; 121 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.52e-14        Length:         702
Score:                  148.00          Matches:        30
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-684-215B-17 (1-30) x ABK14140 (1-702)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 84

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         85 GGGCAGGCGATGGCGATCGCGGGCCAGATC 114

RESULT 26
ABK39769
ID   ABK39769 standard; cDNA; 822 BP.
XX
AC   ABK39769;
XX
DT   21-MAY-2002  (first entry)
XX
DE   DNA encoding lung tumour protein P801P ORF5 and Ra12 fusion protein.
XX
KW   Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW   gene; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200204514-A2.
XX
```

PD 17-JAN-2002.
XX
PF 10-JUL-2001; 2001WO-US022058.
XX
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, Mcnabb A;
PI Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
XX
DR WPI; 2002-164634/21.
DR P-PSDB; AAU85587.
XX
PT Novel polynucleotide encoding a lung tumor polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumor protein.
XX
PS Example 8; SEQ ID NO 1862; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide useful
CC for stimulating and/or expanding T cells specific for a tumour protein
CC for determining the presence of a cancer in a patient. A composition
CC containing the polynucleotide and/or polypeptide is useful for treating a
CC lung cancer in a patient. The polypeptide is useful for removing tumour
CC cells from a biological sample. The polynucleotide is also useful as
CC probe or primer to detect the level of mRNA encoding a tumour protein.
CC This sequence encodes a lung tumour associated protein or protein
CC fragment, described in the method of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.85e-14	Length:	822
Score:	148.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-684-215B-17 (1-30) x ABK39769 (1-822)

```
Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111
```

RESULT 27
ACA12098
ID ACA12098 standard; cDNA; 822 BP.
XX
AC ACA12098;
XX
DT 06-JUN-2003 (first entry)
XX
DE Human lung cancer protein L801P ORF5/Ra12 fusion protein cDNA.
XX
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8; RA12; gene.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US2002197669-A1.
XX
PD 26-DEC-2002.
XX
PF 03-MAY-2001; 2001US-00849626.
XX
PR 13-DEC-2000; 2000US-00736457.
XX
PA (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;
PI Clapper JD;
XX
DR WPI; 2003-352750/33.
DR P-PSDB; ABU69562.
XX
PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT detecting the presence of lung cancer in a patient, and in pharmaceutical
PT compositions, e.g. vaccines, for treating lung cancer.
XX
PS Example 8; Page; 72pp; English.
XX
CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences mentioned in
CC the specification, or a sequence (S2) mentioned in specification,
CC complement of S1, sequences consisting of at least 20 contiguous residues
CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC 90%, identity to S1, or degenerate variants of S1. Also included are an
CC isolated polypeptide (comprising a sequence (S3) selected from any one of
CC the 4 amino acid sequences mentioned in the specification, a sequence
CC encoded by the polynucleotide, or sequences having at least 70%,
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC expression vector comprising the polynucleotide operably linked to an
CC expression control sequence, a host cell transformed or transfected with
CC the vector, an isolated antibody (or its antigen-binding fragment) that
CC specifically binds to the polypeptide, detecting the presence of a cancer
CC in a patient, a fusion protein comprising the polypeptide, an
CC oligonucleotide that hybridises to S1 under moderately stringent
CC conditions, stimulating and/or expanding T cells specific for a tumour
CC protein (comprising contacting T cells with the polynucleotide, protein
CC or antigen-presenting cells, under conditions and for a time sufficient
CC to permit the stimulation and/or expansion of T cells) and inhibiting the
CC development of a cancer in a patient (by incubating CD4^+ and/or CD8^+ T
CC cells isolated from a patient with the polynucleotide, protein or antigen
CC presenting cells that express the polynucleotide, such that T cells
CC proliferate, administering to the patient an effective amount of the
CC proliferated T cells, and thus inhibiting the development of a cancer in
CC the patient. The polynucleotide, protein ands cells are useful in a
CC composition for stimulating an immune response in a patient, and for
CC treating a cancer in a patient (particularly lung cancer). The
CC oligonucleotide is useful for determining the presence of a cancer in a
CC patient. The protein and oligonucleotides are useful in pharmaceutical
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC or primer for nucleic acid hybridisation, and in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and proteins in tumour cells. An amplified portion of the
CC polynucleotide is useful for isolating a full-length gene from a suitable
CC library. The present sequence encodes a fusion protein of human RA12 with
CC the protein product of a cDNA (full length, extended or partial) isolated
CC from a library derived from lung tumour/cancer cells. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?DocId=20020197669
XX

```
SQ   Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.85e-14       Length:         822
Score:                  148.00         Matches:        30
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     7              Gaps:           0

US-09-684-215B-17 (1-30) x ACA12098 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 28
ACA03284
ID   ACA03284 standard; DNA; 822 BP.
XX
AC   ACA03284;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Lung cancer therapyand diagnosis associated DNA #5.
XX
KW   Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ds.
XX
OS   Homo sapiens.
XX
PN   US2002172952-A1.
XX
PD   21-NOV-2002.
XX
PF   10-JUL-2001; 2001US-00902941.
XX
PR   30-JUN-1999;   99US-00346492.
PR   15-OCT-1999;   99US-00419356.
PR   17-DEC-1999;   99US-00466867.
PR   30-DEC-1999;   99US-00476300.
PR   06-MAR-2000; 2000US-00519642.
PR   22-MAR-2000; 2000US-00533077.
PR   10-APR-2000; 2000US-00546259.
PR   27-APR-2000; 2000US-00560406.
PR   05-JUN-2000; 2000US-00589184.
PR   11-JUL-2000; 2000US-00614124.
PR   29-AUG-2000; 2000US-00651563.
PR   08-SEP-2000; 2000US-00658824.
PR   26-SEP-2000; 2000US-00671325.
PR   06-OCT-2000; 2000US-00677419.
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Henderson RA,  Wang T,  Watanabe Y,  Johnson JC,  Retter MW;
PI   Durham M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
XX
DR   WPI; 2003-328427/31.
XX
PT   New polynucleotide, useful for preparing a composition for treating or
PT   inhibiting development of cancer, e.g. lung cancer.
XX
PS   Example 8; SEQID NO 1862; 82pp; English.
XX
CC   The invention describes an isolated polynucleotide comprising one of 32
CC   sequences, complement or degenerate variants of them. The polynucleotide
CC   is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC   for treating or inhibiting development of cancer, e.g. lung cancer. This
CC   sequence represents a polynucleotide associated with the compositions and
CC   methods for the therapy and diagnosis of lung cancer
XX
SQ   Sequence 822 BP; 146 A; 281 C; 270 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.85e-14       Length:         822
Score:                  148.00         Matches:        30
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     7              Gaps:           0

US-09-684-215B-17 (1-30) x ACA03284 (1-822)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 29
ABK39776
ID   ABK39776 standard; cDNA; 861 BP.
XX
AC   ABK39776;
XX
DT   21-MAY-2002  (first entry)
XX
DE   DNA encoding Ral2S-L985P fusion protein.
XX
KW   Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
KW   gene; ss.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200204514-A2.
XX
PD   17-JAN-2002.
XX
PF   10-JUL-2001; 2001WO-US022058.
XX
PR   11-JUL-2000; 2000US-00614124.
PR   29-AUG-2000; 2000US-00651563.
PR   08-SEP-2000; 2000US-00658824.
PR   26-SEP-2000; 2000US-00671325.
PR   06-OCT-2000; 2000US-00677419.
PR   30-OCT-2000; 2000US-00702705.
PR   13-DEC-2000; 2000US-00736457.
PR   03-MAY-2001; 2001US-00849626.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Watanabe Y,  Henderson RA,  Johnson JC,  Retter MW;
PI   Marnerakis M,  Carter D,  Fanger GR,  Vedvick TS,  Bangur CS,  Mcnabb A;
PI   Wang A,  Fanger N,  Switzer A,  Mcneill PD,  Clapper JD;
XX
DR   WPI; 2002-164634/21.
DR   P-PSDB; AAU85594.
XX
PT   Novel polynucleotide encoding a lung tumor polypeptide useful for
PT   stimulating and/or expanding T cells specific for a tumor protein.
XX
PS   Claim 1; SEQ ID NO 1877; 223pp; English.
XX
CC   The invention describes an isolated polynucleotide and polypeptide useful
CC   for stimulating and/or expanding T cells specific for a tumour protein
CC   for determining the presence of a cancer in a patient. A composition
CC   containing the polynucleotide and/or polypeptide is useful for treating a
```

```
CC   lung cancer in a patient. The polypeptide is useful for removing tumour
CC   cells from a biological sample. The polynucleotide is also useful as
CC   probe or primer to detect the level of mRNA encoding a tumour protein.
CC   This sequence encodes a lung tumour associated protein or protein
CC   fragment, described in the method of the invention. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 861 BP; 186 A; 244 C; 209 G; 222 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.96e-14        Length:          861
Score:                    148.00          Matches:         30
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       6               Gaps:            0

US-09-684-215B-17 (1-30) x ABK39776 (1-861)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111

RESULT 30
ACA12105
ID   ACA12105 standard; cDNA; 861 BP.
XX
AC   ACA12105;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Human Ra12S-L985P fusion protein cDNA.
XX
KW   Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW   T cell expansion; CD4; CD8; RA12; gene.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   US2002197669-A1.
XX
PD   26-DEC-2002.
XX
PF   03-MAY-2001; 2001US-00849626.
XX
PR   13-DEC-2000; 2000US-00736457.
XX
PA   (BANG/) BANGUR C S.
PA   (FANG/) FANGER G R.
PA   (WANG/) WANG A.
PA   (WANG/) WANG T.
PA   (SWIT/) SWITZER A P.
PA   (MCNE/) MCNEILL P D.
PA   (CLAP/) CLAPPER J D.
XX
PI   Bangur CS,  Fanger GR,  Wang A,  Wang T,  Switzer AP,  Mcneill PD;
PI   Clapper JD;
XX
DR   WPI; 2003-352750/33.
DR   P-PSDB; ABU69569.
XX
PT   Novel lung cancer polynucleotide encoding lung cancer protein, useful for
PT   detecting the presence of lung cancer in a patient, and in pharmaceutical
PT   compositions, e.g. vaccines, for treating lung cancer.
XX
PS   Example 10; Page; 72pp; English.
XX
CC   The invention relates to a polynucleotide encoding a lung tumour protein,
CC   comprising a sequence selected from any of the 14 sequences mentioned in
CC   the specification, or a sequence (S2) mentioned in specification,
CC   complement of S1, sequences consisting of at least 20 contiguous residues
CC   of S1, sequences that hybridise to S1, sequences having 75%, preferably
CC   90%, identity to S1, or degenerate variants of S1. Also included are an
CC   isolated polypeptide (comprising a sequence (S3) selected from any one of
CC   the 4 amino acid sequences mentioned in the specification, a sequence
CC   encoded by the polynucleotide, or sequences having at least 70%,
CC   preferably 90%, identity to a sequence encoded by the polynucleotide), an
CC   expression vector comprising the polynucleotide operably linked to an
CC   expression control sequence, a host cell transformed or transfected with
CC   the vector, an isolated antibody (or its antigen-binding fragment) that
CC   specifically binds to the polypeptide, detecting the presence of a cancer
CC   in a patient, a fusion protein comprising the polypeptide, an
CC   oligonucleotide that hybridises to S1 under moderately stringent
CC   conditions, stimulating and/or expanding T cells specific for a tumour
CC   protein (comprising contacting T cells with the polynucleotide, protein
CC   or antigen-presenting cells, under conditions and for a time sufficient
CC   to permit the stimulation and/or expansion of T cells) and inhibiting the
CC   development of a cancer in a patient (by incubating CD4^+ and/or CD8^+ T
CC   cells isolated from a patient with the polynucleotide, protein or antigen
CC   presenting cells that express the polynucleotide, such that T cells
CC   proliferate, administering to the patient an effective amount of the
CC   proliferated T cells, and thus inhibiting the development of a cancer in
CC   the patient. The polynucleotide, protein ands cells are useful in a
CC   composition for stimulating an immune response in a patient, and for
CC   treating a cancer in a patient (particularly lung cancer). The
CC   oligonucleotide is useful for determining the presence of a cancer in a
CC   patient. The protein and oligonucleotides are useful in pharmaceutical
CC   compositions, e.g. vaccines. The polynucleotide is also useful as a probe
CC   or primer for nucleic acid hybridisation, and in the design and
CC   preparation of ribozyme molecules for inhibiting expression of tumour
CC   polypeptides and proteins in tumour cells. An amplified portion of the
CC   polynucleotide is useful for isolating a full-length gene from a suitable
CC   library. The present sequence encodes a fusion protein of human RA12 with
CC   the protein product of a cDNA (full length, extended or partial) isolated
CC   from a library derived from lung tumour/cancer cells. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from the USPTO at
CC   seqdata.uspto.gov/sequence.html?DocId=20020197669
XX
SQ   Sequence 861 BP; 186 A; 244 C; 209 G; 222 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.96e-14        Length:          861
Score:                    148.00          Matches:         30
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       7               Gaps:            0

US-09-684-215B-17 (1-30) x ACA12105 (1-861)

Qy          1 ThrAlaAlaSerAspAsnPheGlnLeuSerGlnGlyGlyGlnGlyPheAlaIleProIle 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 ACGGCCGCGTCCGATAACTTCCAGCTGTCCCAGGGTGGGCAGGGATTCGCCATTCCGATC 81

Qy         21 GlyGlnAlaMetAlaIleAlaGlyGlnIle 30
              ||||||||||||||||||||||||||||||
Db         82 GGGCAGGCGATGGCGATCGCGGGCCAGATC 111


Search completed: April 29, 2004, 23:03:07
Job time : 90.3349 secs
```